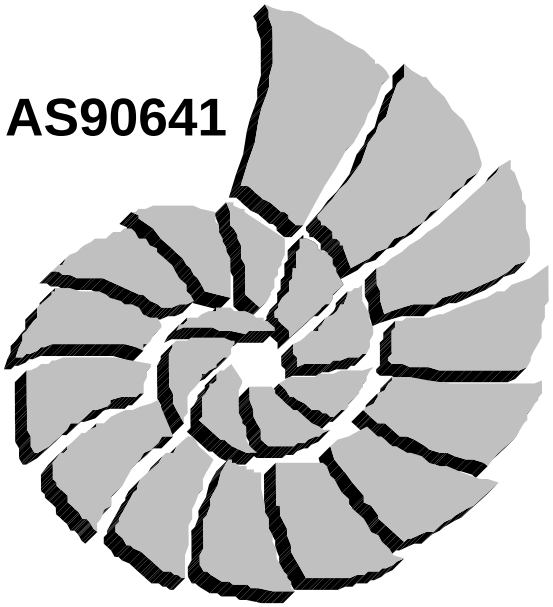


This Is How You Do

Time Series

Kim Freeman

This book covers NZQA, Level 3 Mathematics
Statistics and Modelling 3.1
Determine the Trend for Time Series
Level: 3, Credits: 3, Assessment: Internal

Published by Mahobe Resources (NZ) Ltd

This is How You Do Time Series
Originally published 1999. This edition published 2007 by:
Mahobe Resources (NZ) Ltd
P.O. Box 109-760
Newmarket, Auckland 1149
New Zealand

admin@mahobe.co.nz
www.mahobe.co.nz

ISBN 0-9583564-7-5
ISBN13 9780958356473

'This Is How You Do Time Series' covers the New Zealand Qualifications Authority NCEA Level 3 Mathematics Achievement Standard for Statistics and Modelling 3.1.
The title of the standard is AS90641: Determine the trend for time series data.

Mahobe Resources (NZ) Ltd welcomes feedback on the book and its usefulness.

Contents

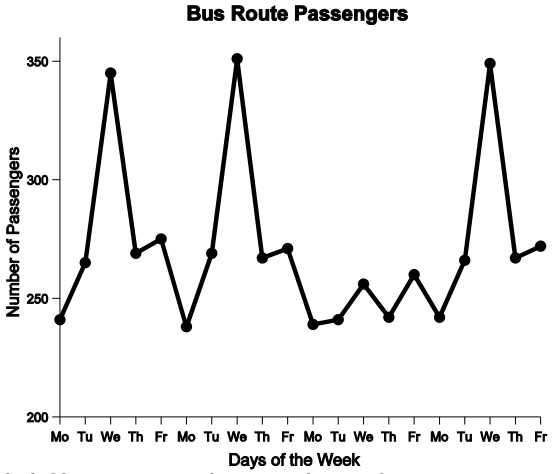
Time Series - Introduction

Time series involves recording data figures at regular intervals over a long period of time. Examples of data could be: retail sales, profits, stock market prices, mortgage rates, currency values, item prices, daily temperatures, weekly rainfall or population growth. Analysing the data over a long period of time, allows you to interpret what is happening and what may happen in the future.

In the table below a bus driver has recorded the number of people travelling on a particular bus route during four weeks in August. Beside the table the data is graphed with days of the week on the horizontal (x) axis and passenger numbers on the vertical (y) axis (starting from 200 passengers).

Question: During August there was one particular week that was very cold and wet. Which of the weeks was it and how can you tell?

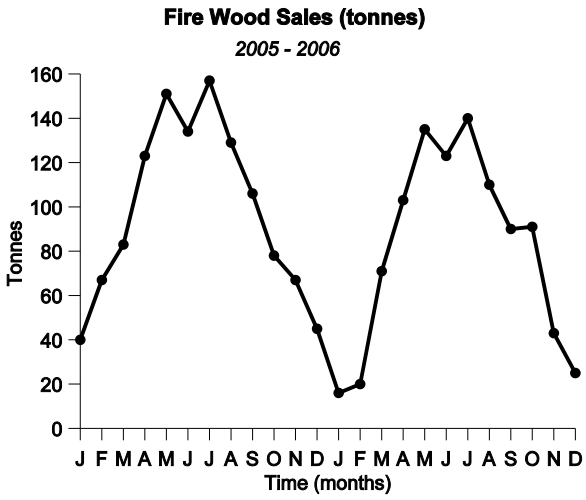
Passenger Numbers					
	Mon	*Tue*	*Wed*	*Thur*	*Fri*
Week 1	241	265	345	269	275
Week 2	238	269	351	267	271
Week 3	239	241	256	242	260
Week 4	242	266	349	267	272



Answer: During the third week there was a marked difference in the number of passengers. This indicates they either stayed at home, or other means of transport.

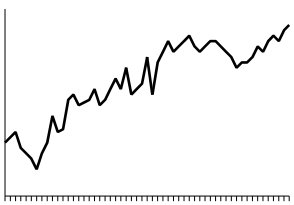
The next table records the monthly amount of firewood sold at a wood merchant's depot. It gives the amounts sold, in tonnes, over a two year period. The graph of the results is also given. The graph is a series of peaks and troughs. Note how over the two years, the peaks and troughs occur at roughly the same time.

Wood Sales *(tonnes)*	*2005*	*2006*
January	40	16
February	67	20
March	83	71
April	123	103
May	151	135
June	134	123
July	157	140
August	129	110
September	106	90
October	78	91
November	67	43
December	45	25

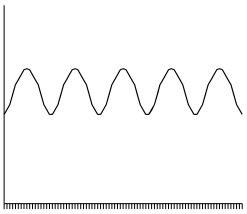


The Use of Times Series

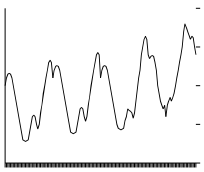
Most time series problems will involve many sets of data and will produce a graph with jagged lines, peaks and troughs. The data enables us to study the past behaviour and therefore predict future events. For example, a clothing store must know when and how many warm jerseys they need to stock and when to order summer wear. By analysing the type and amount of stock sold, the store will be able to purchase the appropriate amount of stock for each season. Below are four types of Time Series graphs that can occur.



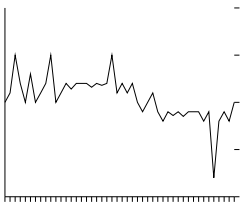
Graph A



Graph B



Graph C



Graph D

Graph A.

Secular Trend - Over a long period of time, data grows or declines. A line of best fit could be placed through these points to illustrate a growth trend. Common secular time series are due to population growth, technological improvements or improvements in business models

Graph B.

Cyclical - Although there are fluctuations, the trend is recurring at regular intervals. In business each cycle is known as prosperity, recession, depression and recovery. Usually each cycle lasts between 2 - 10 years, then repeats itself

Graph C.

Seasonal - These are regular patterns in a business that are not necessarily based around the weather season but can be based around the days of the week, school terms of the year or months of the year.

Graph D.

Irregular - This does not show a regular pattern and the trend is unpredictable. Irregular variations can take place due to earthquake, floods or industrial strikes. A residual effect is when a graph has a sudden dip or spike usually caused by a rare event.

Methods of Measuring Trends in a Time Series Graph

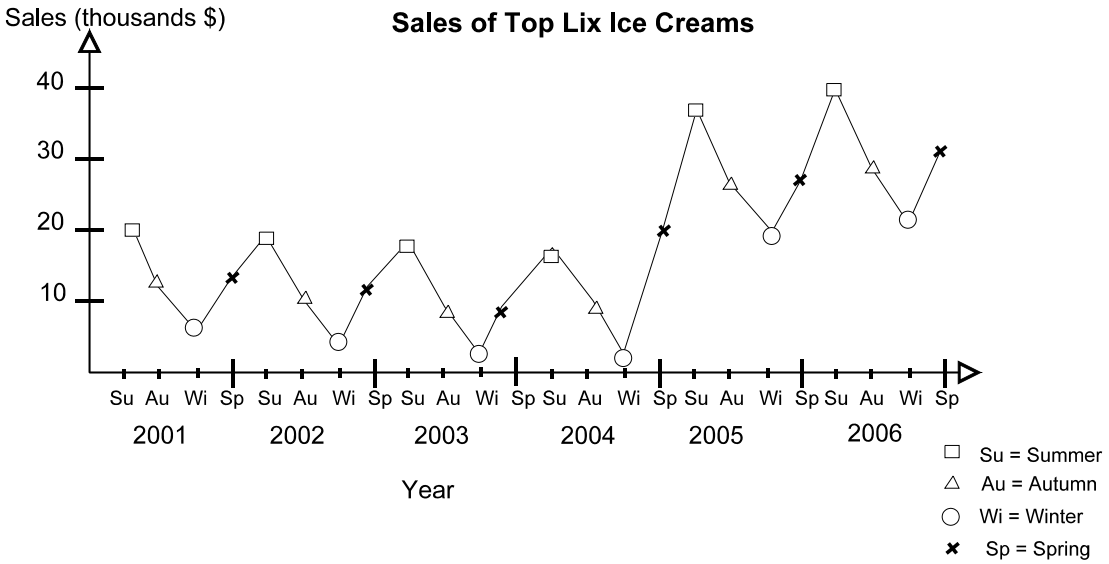
In this eBook we look at 3 methods for measuring the trends in a times series graph:

1. Free Hand Method:
   In this method all of the data is plotted on a graph. A smooth curve is then drawn through the midpoints of each fluctuation. The advantage of this method is that it is simple, flexible and does not need any complex mathematical formula. However its main disadvantage is that it is based on subjective judgements and its lack of mathematical accuracy can lead to bias of the results.

2. Moving Means Method
   This method smooths out seasonal variations in a graph by taking data averages over each cycle. When calculating moving means, take the same number of intervals as the length of the seasonal patterns. The advantage of this method is that it is easy and simple to compute. The disadvantage is that if the proper period of the moving means is not used then the results can be misleading.

3. Least Square Method
   Mathematically this is the most accurate method of finding a trend line. This approach can be used to fit a straight line, parabolic trend or exponential trend. In this book we will only deal with the straight line trend. The calculations used in this method can be quite time consuming, however a number of graphic calculators and/or Excel can easily compute the equations of a straight line, parabolic or exponential trend. The method involves taking the sum of the deviations from the actual values and forming a mathematical equation which can be used for forecasting. The disadvantage is that the computations can be complex and if data is added later then all computations have to be repeated.

Freehand Trend Lines - Exercises



1. Mahobe own an ice-cream van and a business called Top Lix. They keep a record of sales over 6 years. During 2005, Mahobe increased their sales area and working hours in an effort to get more business. The graph above uses the data from all sales records.

   a. Identify at least 3 features of the graph.

   . . . . . .

   . . . . . .

   . . . . . .

   . . . . . .

   . . . . . .

   . . . . . .

   b. How could you show the overall trend of the sales?

   . . . . . .

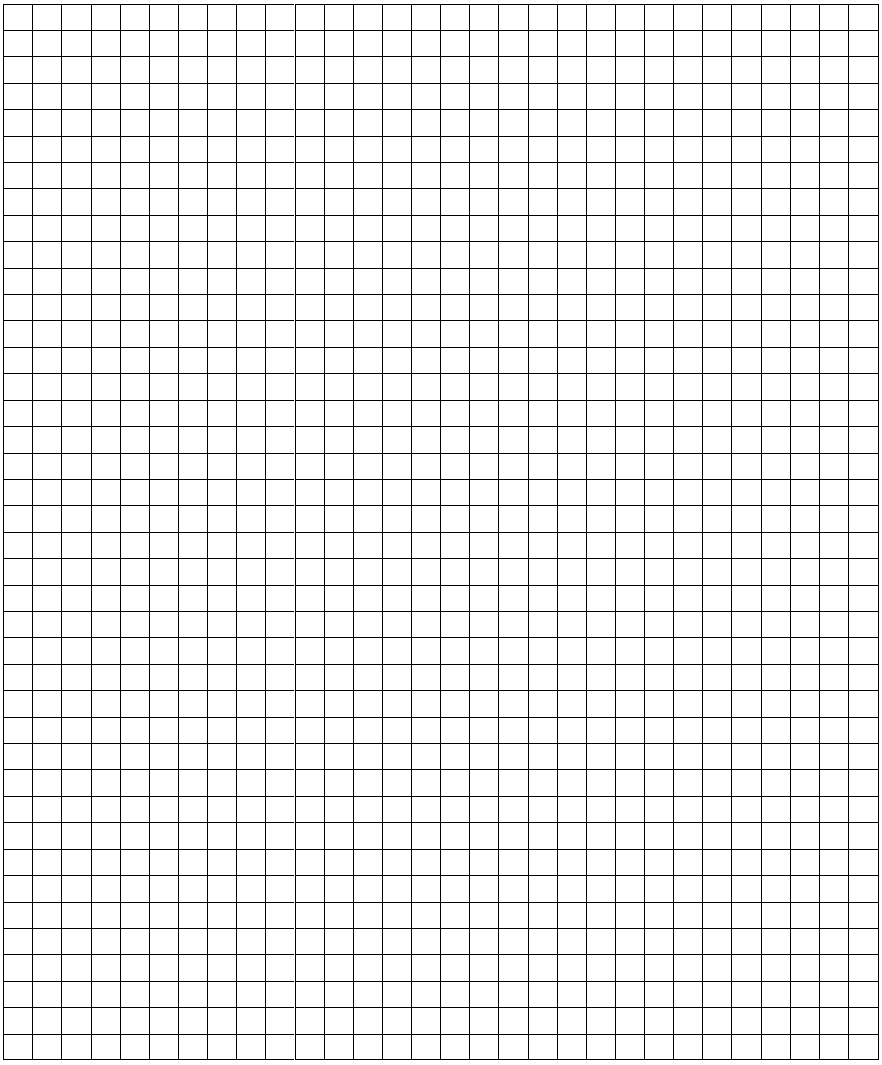
   . . . . . .

   . . . . . .

2. Graph the following data, then draw a freehand trend line that shows what has happened to sales over the 5 years.

Mahobe Takeaway
Quarterly Profit

Sept 2001	6254
Dec 2001	6953
March 2002	6381
June 2002	6350
Sept 2002	6290
Dec 2002	7090
March 2003	6189
June 2003	6152
Sept 2003	6167
Dec 2003	6799
March 2004	6351
June 2004	6295
Sept 2004	6426
Dec 2004	7302
March 2005	6634
June 2005	6698
Sept 2005	6843
Dec 2005	7732
March 2006	7097
June 2006	7251
Sept 2006	7423



Describe any features that you notice.

. . . . . .

. . . . . .

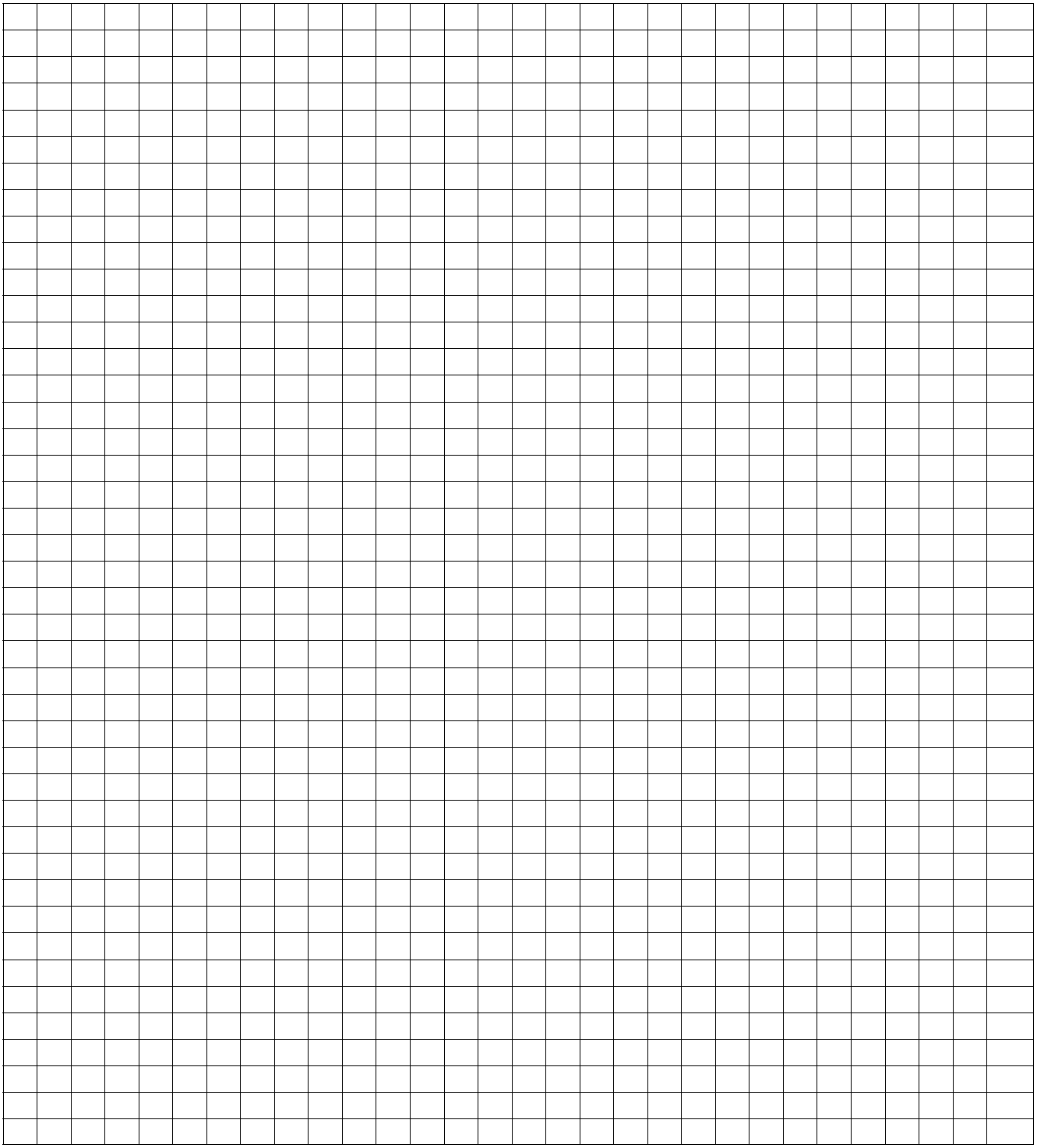
. . . . . .

. . . . . .

3. Graph the following data and then draw a freehand trend line. Describe what is happening to the data.

   Sales of Mathematics texts - Mahobe Resources.

Year	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Sales ('000 units)	25	32	30	31	33	30	34	35	37	18	36



4. Graph the following data and then draw a freehand trend line. Describe what is happening to the data.

Production of Cars - Mahobe Car Plant 2006

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
No. Of Cars Produced	320	267	354	320	270	245	270	281	315	325	351	306

Moving Means

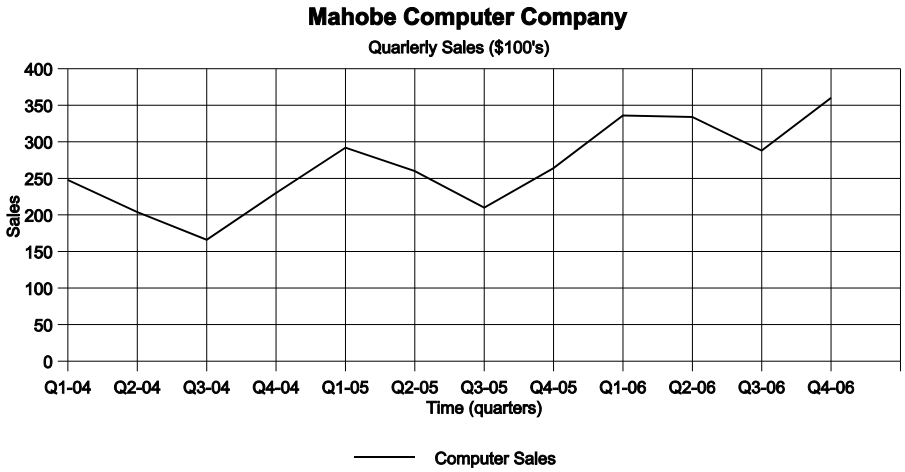
Because trends can be mixed with seasonal variations there is a method used to smooth out the graph. It is called the method of moving means. When calculating moving means, take the same number of intervals as the length of the seasonal patterns. e.g. If the data are daily, average them out for the week by taking 7 intervals. If the data are monthly, average them for the year by taking 12 intervals and if the data are quarterly, average them by taking 4 intervals.

e.g. Mahobe Computer Company records the sales of computers over a three year period. The sales are given below. A graph of the sales data is also given below.

By using the method of moving means find the trend, then forecast sales for the first two quarters of the year 2007.

Sales by Mahobe Computer Company ($100's)
Quarters

	1	*2*	*3*	*4*
2004	248	204	166	230
2005	292	260	210	264
2006	336	334	288	360



Note: In some Mathematics texts the term “Moving Mean” is used while in others the term “Moving Average” is used. They are the same thing.

Calculating Moving Means

Table B

Table of Moving Means and Seasonal Variation

1 Year	2 Quarter	3 Sales	4 Yearly Total	5 Moving Mean	6 Centred Mean (Trend Line)	7 Seasonal Variation
2004	1	248				
	2	204				
			848	212		
	3	166			218	-52
			892	223		
	4	230			230	0
			948	237		
2005	1	292			243	49
			992	248		
	2	260			252	8
			1026	257		
	3	210			262	-52
			1070	268		
	4	264			277	-13
			1144	286		
2006	1	336			296	40
			1222	306		
	2	334			318	16
			1318	330		
	3	288				
	4	360				

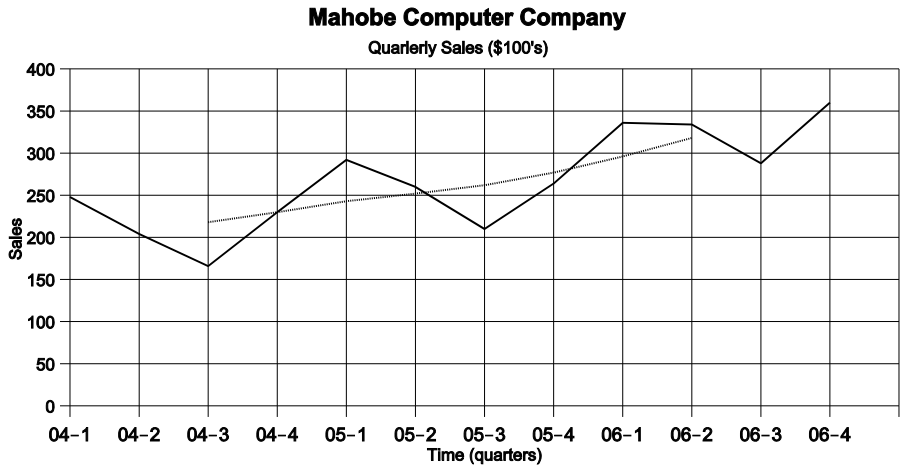
Column 4. The moving sum of the computer sales during four successive quarters. The numbers are obtained by adding each successive four quarters. The sum is placed in the middle of the numbers being totalled.

Column 5. The moving means for consecutive 4 quarter periods. These are obtained by dividing each yearly total (column 4) by 4. These figures represent the mean quarterly sales of computers during the year. The seasonal fluctuations are averaged out.

Column 6. The mean of the two nearest moving means. This allows the centred mean to correspond to that time when the sales occurred. These figures (called the centred mean) are plotted to become the trend line.

Column 7. Contains the difference between quarterly sales and the moving mean (trend) for that quarter. Note, if column six was not "centred", we could not have done this.

Graph B



Graph B shows the trend line (the centred mean) added to the original graph. Note how all seasonal fluctuations have been averaged out.

Calculating Seasonal Factors

Each quarter has a seasonal factor associated with it. Because all of the factors in column 7 (Table B) are different, it is helpful to devise another table. In Table C, below, the total row consists of the sums of the variations (from column 7) and the mean (last) row contains their averages. This mean row gives the four seasonal factors. These factors can be subtracted from any corresponding quarter to deseasonalise the quarterly value. Alternatively it can be added to the trend line value to predict a seasonalised quarterly value.

Table C

	Quarter 1 Jan, Feb, Mar	Quarter 2 Apr, May, Jun	Quarter 3 Jul, Aug, Sep	Quarter 4 Oct, Nov, Dec
2004			-52	0
2005	49	8	-52	-13
2006	40	16		
Total	89	24	-104	-13
Mean	44.5	12	-52	-6.5

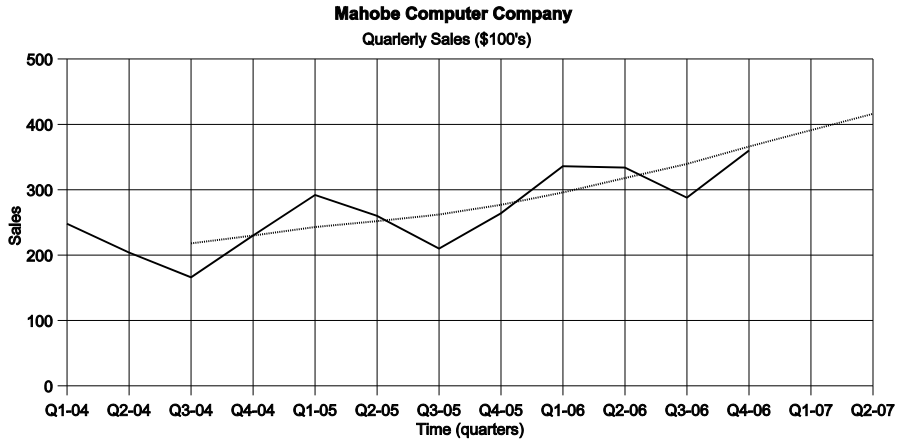
Extending the Trend Line

The trend line and the seasonal factors can now be used to predict the anticipated sales for the next two quarters. Third quarter 2006 sales were 288 ($100's). The 3rd quarter seasonal factor is -52. This means -52 is subtracted from 288 (in order to deseasonalise). The same applies to the 4th quarter data of 360 ($100's) from which -6.5 is subtracted.

Third Quarter 2006 288 – -52 = 340
Fourth Quarter 2006 360 – -6.5 = 366.5

Graph C, below, shows how this data is used to extend the trend line. Using these figures, the trend line can be projected by two more quarters.

Graph C



Note how the seasonal factors in Table C work. These give the mean of each quarterly value: 44.5, 12, -52 and -6.5 (i.e. the seasonal variation). These figures indicate that (on average) Quarter 1 will be 44.5 above the trend line, Quarter 2 will be 12 above the trend line, Quarter 3 will be 52 below the trend line and Quarter 4 will be 6.5 below the trend line. Using the projected trend line, Mahobe Computer Company can reasonably expect to have over $39,000 worth of sales in the first quarter, and over $41,000 in the second quarter. This is because historically Q1 and Q2 are regularly above the trend line. According to the sales figures over the last three years, Mahobe should look forward to increased sales over the first half year (assuming the trend continues).

Moving Means - Exercises

1. The table below gives the sales of televisions sold at the Mahobe Vision Shed. Column B gives the 3 point moving mean of Column A. Column C gives the 3 point moving mean of Column B.

	A - Monthly Sales	B - Moving Mean (3 point)	C - Moving Mean (3 point)
Apr	10		
May	36	31.00	
Jun	47	31.33	31.78
Jul	11	**A**	33.33
Aug	41	35.67	36.11
Sep	55	39.67	40.00
Oct	23	44.67	**B**
Nov	56	42.67	
Dec	49		

   a. Calculate the values of A and B. Write the values into the table.

   b. How many Monthly Sales values are involved in obtaining the value 40.00 in Column C?

   . . . . . .

2. The table below gives sales from the 'Mahobe Cola Company'. It shows the amount of Cola sold each quarter over a 3 year period. Calculate the missing values then draw a graph of the sales data and the centred moving mean data. Comment on the graph.

Quarter	Millions of Litres	Mean (4 Point Moving Mean)	Centred Mean (Centred Moving Mean)
Sept 2003	13.0		
Dec 2003	18.5	13.73	
March 2004	11.1	13.43	13.58
June 2004	12.3	13.48	13.45
Sept 2004	11.8	**A .....**	13.46
Dec 2004	18.7	13.10	13.28
March 2005	11.0	13.45	13.28
June 2005	10.9	13.53	13.49
Sept 2005	13.2	13.50	13.51
Dec 2005	19.0	14.03	13.76
March 2006	10.9	13.75	**B .....**
June 2006	13.0		
Sept 2006	12.1		

3. The data below shows New Zealand's production of goods and services relative to a base year.

Year	1999/2000	2000/2001	2001/2002	2002/2003	2003/2004	2004/2005	2005/2006
Index of Gross Domestic Product	1056	1103	1182	1230	1251	1252	1201

a. Graph the series above along with a three year moving mean.

b. List the components generally regarded as combining to form time series data, then use the figures above, to help explain how each of the relevant components affects this series.

. . . . . .

. . . . . .

. . . . . .

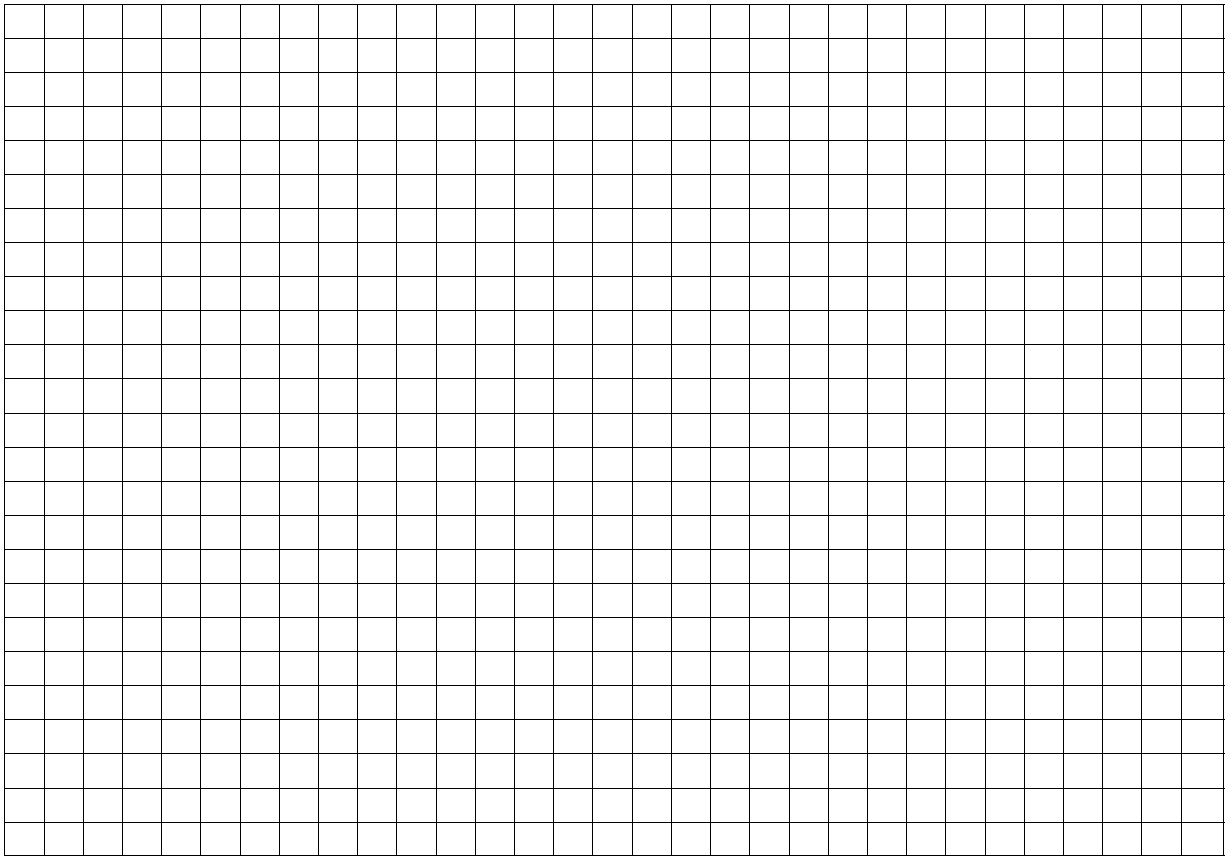
. . . . . .

. . . . . .

. . . . . .

4. The following data represents quarterly shoe sales for the Mahobe Sports Shoe Company.
   a. Use the grid below to draw a graph of the data.

Year	Quarter	Sales ($ million)
2003	1	3.6
	2	4.4
	3	4.5
	4	10.6
2004	1	3.8
	2	4.6
	3	4.7
	4	11.2
2005	1	4.2
	2	5.0
	3	5.1
	4	11.8
2006	1	4.2
	2	5.0
	3	5.1
	4	11.8



   b. What would happen if you simply drew a graph of the figures above and then put in a "line of best fit"? Could you reliably make predictions based on this graph?

. . . . . .

. . . . . .

. . . . . .

. . . . . .

c. Complete the next two tables using the figures given. Use your calculations to graph (on the previous page) the centred moving mean trend line, and predict sales for the first two quarters of the year 2007.

Year	Quarter	Sales
2003	1	3.6
	2	4.4
	3	4.5
	4	10.6
2004	1	3.8
	2	4.6
	3	4.7
	4	11.2
2005	1	4.2
	2	4.9
	3	4.8
	4	11.8
2006	1	4.2
	2	5.0
	3	5.1
	4	11.8

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2003				
2004				
2005				
2006				
Total				
Mean				
Seasonal Factors				

Conclusions: . . . . . .

. . . . . .

. . . . . .

5. The time series in the table below gives the numbers of registered unemployed in Mahobe City on 1 April, 1 August and 1 December over the last five years. The means of each successive three term set and individual four-month effects are included with one value missing in each of the two columns.

Time		Number	Mean of Three	Individual Seasonal Effects
2002	July 31	3851		
	November 31	3412	3528	-116
2003	March 31	3321	3453	-132
	July 31	3626	3360	266
	November 31	3133	3271	-138
2004	March 31	3054	3214	-160
	July 31	3455		
	November 31	2869	3033	-164
2005	March 31	2775	2934	-159
	July 31	3158	2847	311
	November 31	2608	2751	-143
2006	March 31	2487	2658	-171
	July 31	2879		

a. Calculate the missing mean and the missing seasonal effect for July 31, 2004.

. . . . . .

b. Use the individual seasonal effects to obtain estimates for the seasonal effects for July, November and March.

. . . . . .

. . . . . .

. . . . . .

c. Seasonally adjust the number unemployed on July 31, 2006. Comment on a claim by a journalist in the local newspaper that unemployment in the city is rising.

. . . . . .

. . . . . .

d. Assume that the trend line continues to change with an extra 65 people registering as unemployed at the end of each successive 4-month interval. Use the estimates of the seasonal effects from b. to forecast unemployment levels for March and July 31 2007.

. . . . . .

. . . . . .

. . . . . .

6. Quarterly sales data for 2004 - 2006 from Mahobe Clothing and Toy Store are given below. Sales are in thousands ($000). The table is incomplete.

Time	Sales ($000)	Centred Moving Mean (order 4)	Individual Seasonal Effects
1. March 2004	352		
2. June 2004	438		
3. September 2004	388	414.125	-26.13
4. December 2004	461	420.125	40.88
5. March 2005	387	426.375	-39.38
6. June 2005	451	436.000	15.00
7. September 2005	425	445.875	
8. December 2005	501	453.625	47.38
9. March 2006	426	459.375	-33.38
10. June 2006	474	465.750	8.25
11. September 2006	448		
12. December 2006	529		

a. Find the individual seasonal effect for September 2005.

. . . . . .

b. Find the mean seasonal effect for the June quarter using these data.

. . . . . .

c. Which quarters are used to calculate the centred moving mean for the June 2005 quarter of 436.00?

. . . . . .

d. The seasonal effect for the March quarter is -36.38.
   i. If the deseasonalised sales for the March quarter of 2007 were $478 000, calculate the actual sales for the quarter.

   . . . . . .

   ii. The trend in sales is approximately linear and can be described by the equation:
   Sales = 389 + 6.5Q where Q is the number of quarters after the December 2003 quarter.
   Use both the trend line and the March seasonal effect to forecast sales in the March quarter of 2008.

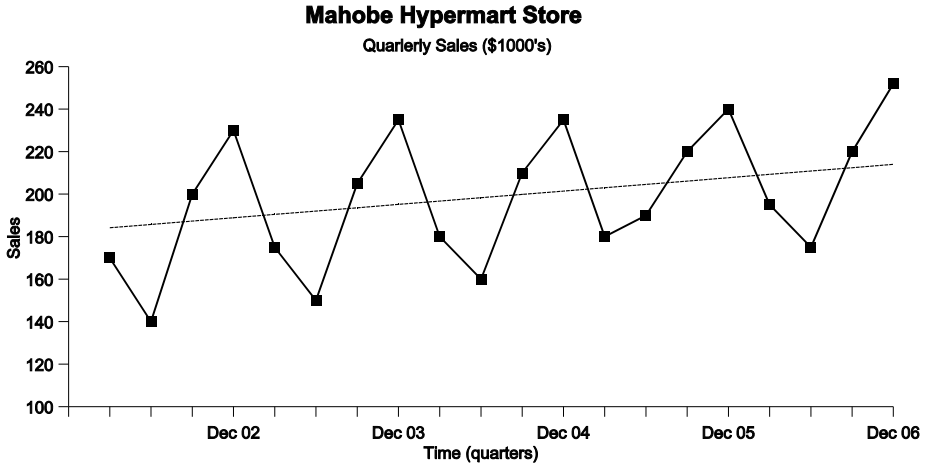
   . . . . . .

   . . . . . .

   . . . . . .

   . . . . . .

7. Quarterly sales figures (in $1000s) for the Mahobe Hypermart Store for the period March 2002 (t = 1) to December 2006 (t = 20) are shown in the time series plot below. The marketing department have also produced a trend line (T) that is given by the equation T = 182.6 + 1.57t.



a. Desc
ribe the features of the graph that show it has seasonal variation.

. . . . . .

. . . . . .

b. Using a centred moving mean, the quarterly effects have been calculated and given in this table.

Quarter	March	June	Sept	December
Quarterly Effect	-16.88	-31.44	12.97	37.34

i. Give a possible value for the number of points (n) when calculating the centred moving mean.

. . . . . .

ii. Explain why in this situation a centred moving mean is used rather than a moving mean.

. . . . . .

. . . . . .

iii. Using the model above, forecast sales (in $1000s) for each of the quarters during 2007.

. . . . . .

. . . . . .

. . . . . .

. . . . . .

. . . . . .

8. Investment in Mahobe Shares on the stock market have been tracked for each quarter. The table below shows the value of stock being purchased (in $millions) for the period March 2003 until June 2006.

Quarter	Stock Value ($millions)	Centred Moving Mean	Average Seasonal Factors	Seasonally Adjusted Value
Sep 03	6100	6636	-229	6329
Dec 03	6650	6585	58	**
Mar 04	6700	6655	46	6654
Jun 04	6830	6813	-26	6856
Sep 04	6750	6976	-229	6979
Dec 04	7230	7158	58	7172
Mar 05	7430	7383	46	7384
Jun 05	7550	7619	-26	7576
Sep 05	7860	7814	**	8089
Dec 05	8040	8004	58	7982
Mar 06	8180		46	8134
Jun 06	8320		-26	8346

a. Calculate i. The Seasonal Index for September 2005. . . . . .

ii. The Seasonally Adjusted Turnover values for December 2003. . . . . .

b. If the figures for Stock Value, Centred Moving Mean and Seasonally Adjusted Value were graphed, which line would be termed the "Trend Line"?

. . . . . .

c. Was the June 2005 quarter higher or lower than expected? Justify your answer with reference to the information in the table.

. . . . . .

. . . . . .

. . . . . .

. . . . . .

d. If the centred moving mean increases on average by 205 each quarter, use the values in the table to forecast turnover in December 2006.

. . . . . .

. . . . . .

. . . . . .

9. The occupancy rate for Mahobe Hotel for each of the four seasons over the last four years is shown in the table below. Also included in the table are: Moving Mean (the mean of four successive terms in the time series for the motel's occupancy rate), Centred Moving Mean (the mean of two successive moving means) and Individual Seasonal Effects (the difference between the occupancy rate and the centred moving mean). One value is missing from each of the last three columns.

Season	Occupancy Rate	Moving Mean	Centred Moving Mean	Individual Seasonal Effects
2002 Winter	0.664			
Spring	0.892	0.7690		
2003 Summer	0.875	0.7610	**	0.1100
Autumn	0.645	0.7588	0.7599	-0.1149
Winter	0.632	0.7560	0.7574	-0.1254
Spring	0.883	0.7388	0.7474	**
2004 Summer	0.864	0.7143	0.7265	0.1375
Autumn	0.576	0.6885	0.7014	-0.1254
Winter	0.534	0.6493	0.6689	-0.1349
Spring	0.780	0.6453	0.6473	0.1328
2005 Summer	0.707	0.6325	0.6389	0.0681
Autumn	0.560	0.6143	0.6234	-0.0634
Winter	0.483	**	0.6104	-0.1274
Spring	0.707	0.5920	0.5993	0.1078
2006 Summer	0.676			
Autumn	0.502			

a. Calculate the 3 missing values in the table.

. . . . . .

b. Use the individual Summer seasonal effects to estimate the seasonal effect for Summer.

. . . . . .

. . . . . .

. . . . . .

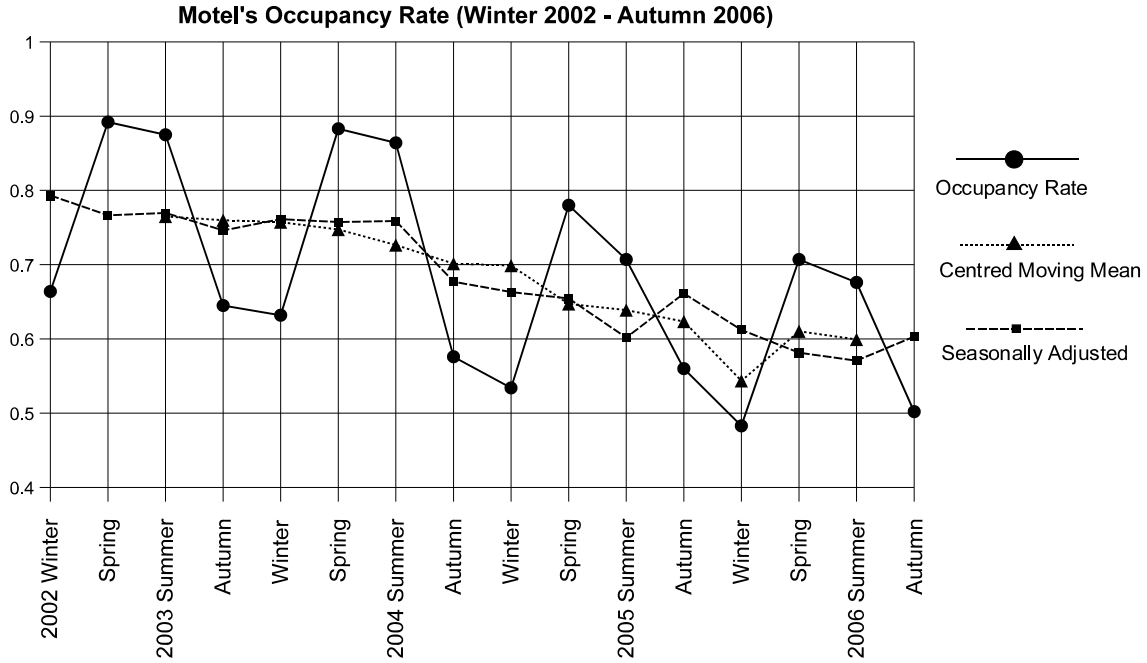
c. Seasonally adjust the occupancy rate for Summer 2005. Was it higher or lower than expected?

. . . . . .

. . . . . .

. . . . . .

The occupancy rates, the centred moving means and the seasonally adjusted occupancy rates for the motel are shown in the plot below.



d. What do the centred moving means show us about the motel's occupancy rates over this period of time?

. . . . . .

. . . . . .

. . . . . .

e. Study the occupancy rates for 2005. For each of the Summer, Winter, Autumn and Spring plots say whether the occupancy was better or worse than expected. Justify each of your answers by using information shown in the graph above.

. . . . . .

. . . . . .

. . . . . .

f. Assuming that the centred moving means for the motel's occupancy rate changes by 0.01 for each season after 2006, use the estimate of the seasonal effect for Summer in b. to forecast the motel's occupancy rate for Summer 2007.

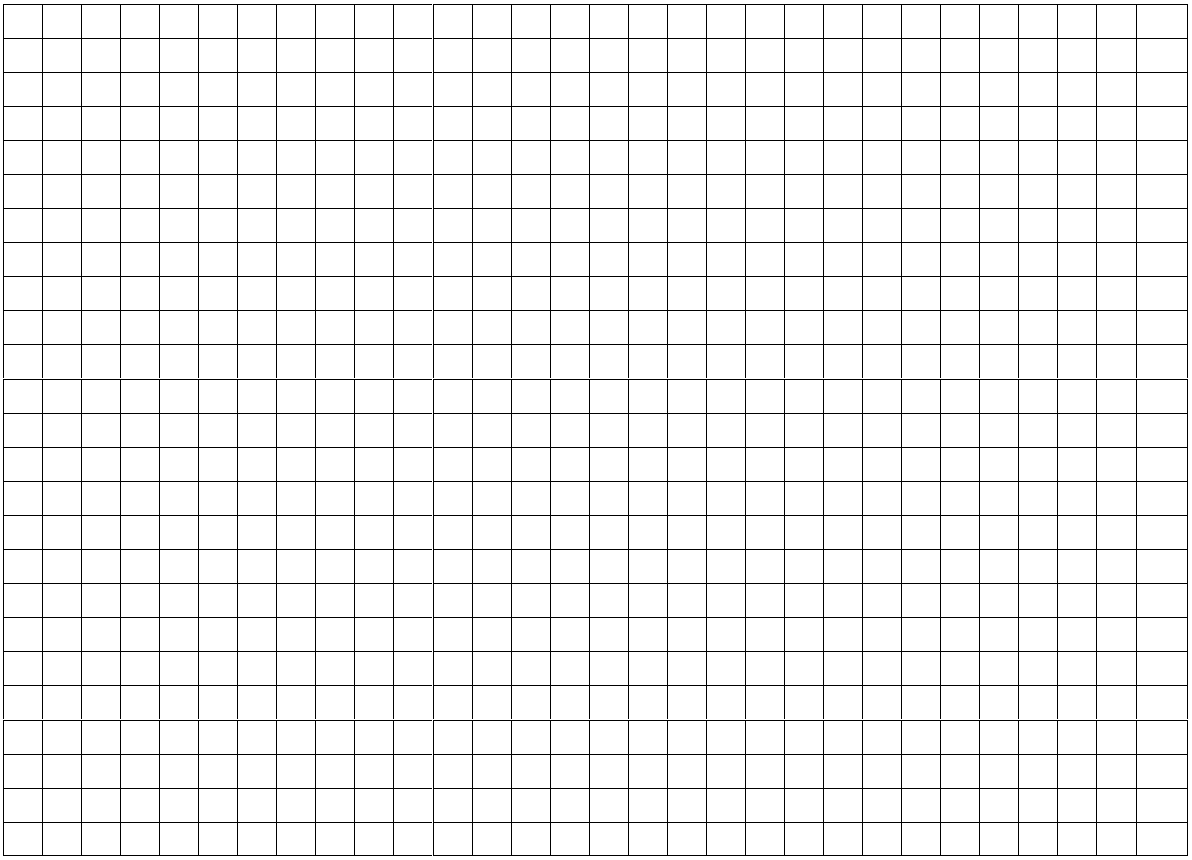
. . . . . .

. . . . . .

10. The table below gives the units of electricity used by a household for a period of 19 weeks.

Week	Units Used	Moving Means	Week	Units Used	Moving Means
1	65		**11**	74	
2	73		**12**	75	
3	72		**13**	73	
4	68		**14**	67	
5	73		**15**	67	
6	86		**16**	62	
7	79		**17**	60	
8	87		**18**	62	
9	82		**19**	62	
10	82				

Plot the data and on the same graph plot a 3 point and a 5 point moving mean. What does the moving mean line indicate? Discuss the differences between the two moving means. Are the figures above useful in identifying seasonal variations in electricity usage by the household?



. . . . . .

. . . . . .

. . . . . .

. . . . . .

. . . . . .

Least Squares Method

This is a mathematical method for time series trend-fitting and is the method used most often in real-life practice. The method results in a mathematical equation which can be used in forecasting future trend values. The disadvantage of this method is that the computations can be seen as complex and if future figures are added then all computations have to be redone. Computations are best done using a spreadsheet.

e.g. The following data gives the yearly profit (in $000s) for Mahobe Airways since 2000. Fit a straight line trend line using the method of least squares (taking the year 1999 as the year of origin). Estimate the values for 2007 and 2010.

Year	1999	2000	2001	2002	2003	2004	2005	2006
Profit ($000s)	125	128	133	135	140	141	143	147

Year	**Profit (Y)**	**Deviation from 1999 (X)**	**XY**	**X^2**
1999	125	0	0	0
2000	128	1	128	1
2001	133	2	266	4
2002	135	3	405	9
2003	140	4	560	16
2004	141	5	705	25
2005	143	6	858	36
2006	147	7	1029	49
N = 8	$\sum Y = 1092$	$\sum X = 28$	$\sum XY = 3951$	$\sum X^2 = 140$

A straight line can be expressed by the equation: $Y = a + bX$
where: Y = Trend Value, X = the unit of time, a = the y-intercept, b = the slope of the line

The normal equations are: $\sum Y = Na + b\sum X$
and $\sum XY = a \sum X + b \sum X^2$

Substituting values: $1092 = 8a + 28b$
$3951 = 28a + 140b$

Multiplying equation 1 by 7: $7644 = 56a + 196b$
Multiplying equation 1 by 2: $7902 = 56a + 280b$
Subtracting $-258 = -84b$
$b = 3.071$

Substituting b into and equation 1 $1092 = 8a + 28(3.071)$
$a = 125.752$

Therefore the trend equation is $Y = 125.752 + 3.071X$

At 2007, $X = 8 \therefore 125.752 + 3.071(8) = 150.32$, i.e. $150 320 profit

At 2010, $X = 11 \therefore 125.752 + 3.071(11) = 159.533$, i.e. $159 533 profit

An alternate method is to find the slope of the trend by the formula: $b = \dfrac{\sum xy - \frac{(\sum y)(\sum x)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$

and the intercept of the trend by the formula: $a = \dfrac{\sum y}{n} - b\left(\dfrac{\sum x}{n}\right)$

In the next example the origin for least squares is taken from the middle of the data. This allows the trend formula to be substantially reduced. Note that for this method, there needs to be an odd number of data.

Year	1999	2000	2001	2002	2003
Sales ($00s)	225	280	390	230	375

Year	**Profit (Y)**	**Deviation from 2001 (X)**	**XY**	**X^2**
1999	225	-2	-450	4
2000	280	-1	-280	1
2001	390	0	0	0
2002	230	1	230	1
2003	375	2	750	4
N = 5	**$\sum Y$ = 1500**	**$\sum X$ = 0**	**$\sum XY$ = 250**	**$\sum X^2$ = 10**

Because $\sum X = 0$, both formula can be reduced to: $a = \dfrac{\sum y}{n}$ and $b = \dfrac{\sum xy}{\sum x^2}$

Substituting values we get: $a = \dfrac{1500}{5}$ and $b = \dfrac{250}{10}$ $\therefore$ the trend line = Y = 300 + 25X

The least squares method is also helpful if there are gaps in the data. In the example below the least squares method works around the lack of data for 2001.

Year	1999	2000	2002	2003	2004
Sales ($00s)	154	176	188	190	182

Year	**Sales (Y)**	**Deviation from 1999 (X)**	**XY**	**X^2**
1999	154	0	0	0
2000	176	1	176	1
2002	188	3	564	9
2003	190	4	760	16
2004	182	5	910	25
N = 5	**$\sum Y$ = 890**	**$\sum X$ = 13**	**$\sum XY$ = 2410**	**$\sum X^2$ = 51**

There are many more examples that could be given such as how to use the above quick formulae with an even set of data values or how to fit a parabola or exponential trend to a given set of values. Also, with the 'Chart' option on a spreadsheet you can also 'add trend line' and display the equation of a trend line no matter whether the line formed by the data is linear, logarithmic, polynomial, power, exponential or a moving average.

The use and understanding of Moving Means is sufficient to gain excellence in this course. The important part of your internal assessment is answering all the required questions in your report.

Least Squares Method - Exercises

1. Fit a straight line trend by using the method of least squares (taking 1999 as the year of origin) to the following data. Calculate the trend values for each year and compare them to the actual values.

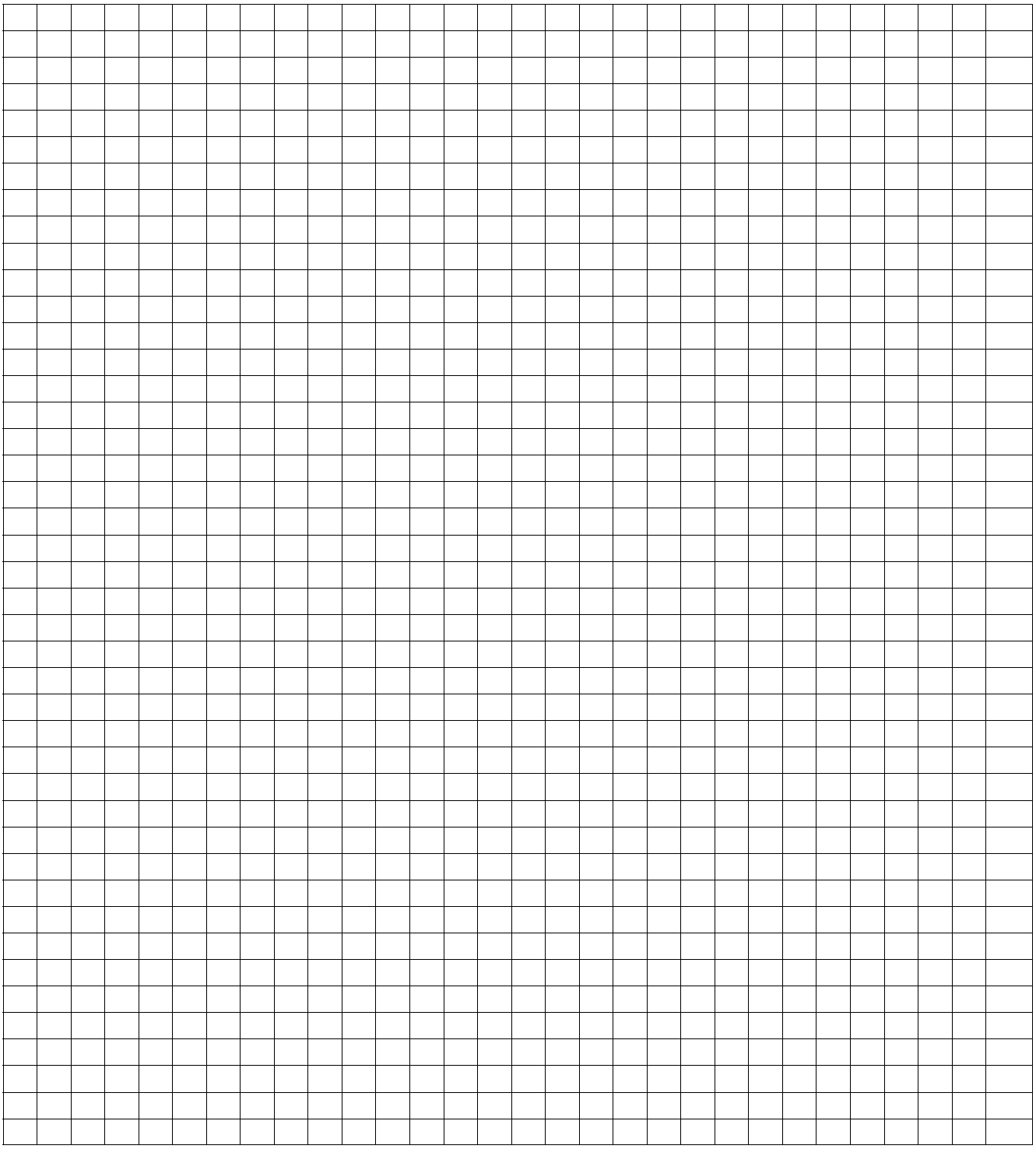
Year	2000	2001	2002	2003	2004	2005
Value	10	15	21	25	32	40

Year	**Value (Y)**	**Deviation from 1999 (X)**	**XY**	**X^2**
2000	10	1	10	1
2001	15	2	30	4
2002	21	3	63	9
2003	25	4	100	16
2004	32	5	160	25
2005	40	6	240	36
N =	$\sum Y$ =	$\sum X$ =	$\sum XY$ =	$\sum X^2$ =

Year	2000	2001	2002	2003	2004	2005
Actual Value	10	15	21	25	32	40
Trend Values						

2. The data below gives sales in $millions for Mahobe Oil Company. Draw a graph of the data and fit a straight line trend by using the method of least square (taking 1987 as the year of origin). Comment on the trend line.

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
Sales	2.8	3.0	3.5	4.0	4.6	5.0	5.4	6.0	7.0	8.0
Year	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Sales	9.7	10.3	10.8	10.2	10.6	10.6	11.5	13.3	17.0	18.4

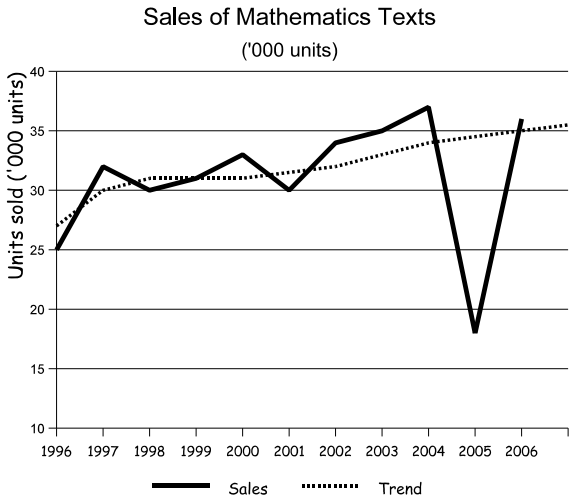


The Answers

Page 4

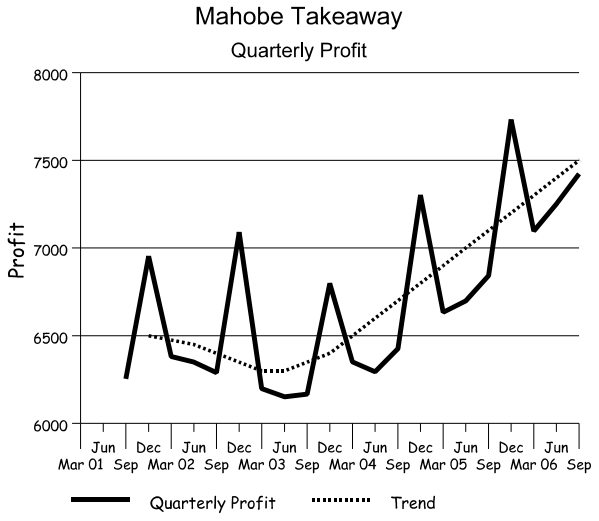
1. **a.** **i.** The graph is seasonal with sales having definite high medium and low periods. In this example the sales seem to be based around the yearly seasonal cycles.
ii. From 2001 to 2004 sales were gradually declining. To see this, draw a freehand line through each of the season's corresponding points and notice the line's negative gradient.
iii. Something happened after the winter of 2004 as sales dramatically increased. Sales increase each season (again draw a freehand line through each seasons points and notice that the gradients are positive). However note that there is only limited data to support this trend.

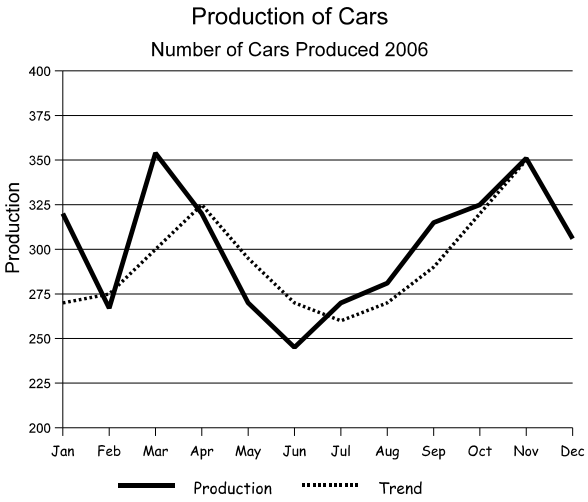
b. As noted in iii. (above) you could draw freehand lines through each of the seasonal points to show the trend or to predict future sales.



Page 5

2. Mahobe Takeaway: The time series data has a quarterly seasonal pattern (increased sales between September & December). There was a slight decrease in sales for the first two years but since then profits have increased. The graph of the data and a freehand trend is below. Note that with this sort of graph it is hard to draw a freehand trend as the December seasonal high causes an overestimation of the other quarters.





Page 6 - 7

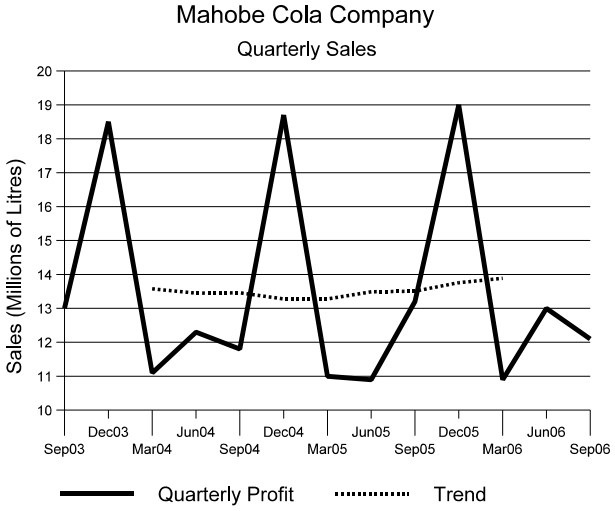
3. Overall there are increasing sales however it is very hard to predict as there is at least one random component that has affected sales.

4. Shows seasonal components with major peaks in March, November and a minor peak in January. Traditionally car sales peak just before the end of the financial year (March 31), just before Christmas and in the January holidays. However because there is only 1 year's worth of data we cannot compare seasonal periods and show whether there is a general growing or declining in production.

Page 11

1. A: (47 + 11 + 41) ÷ 3 = 33.0
B: (39.67 + 44.67 + 42.67) ÷ 3 = 42.34
There are 5 monthly sales involved: 11, 41, 55, 23, and 56. It also involves 4 calculations.

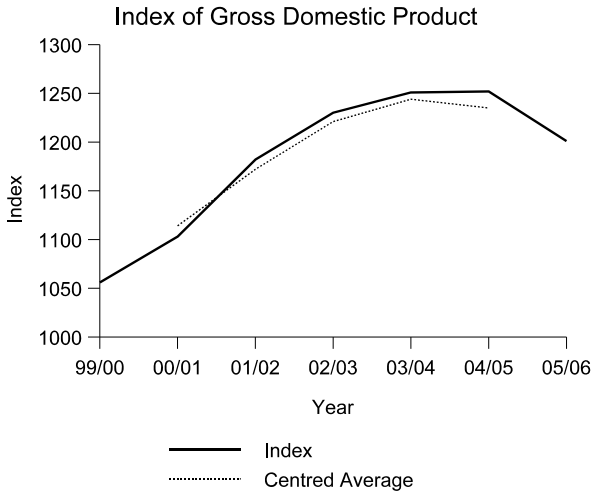
2. Mahobe Cola Company - graph is below
A = 13.45, B = 13.89
The time series data has a quarterly seasonal pattern (increased sales between September & December). The trend line (moving mean) indicates that since March 2005 there has been a slight but steady increase in sales.



Page 12

3. a. Data for the graph is:

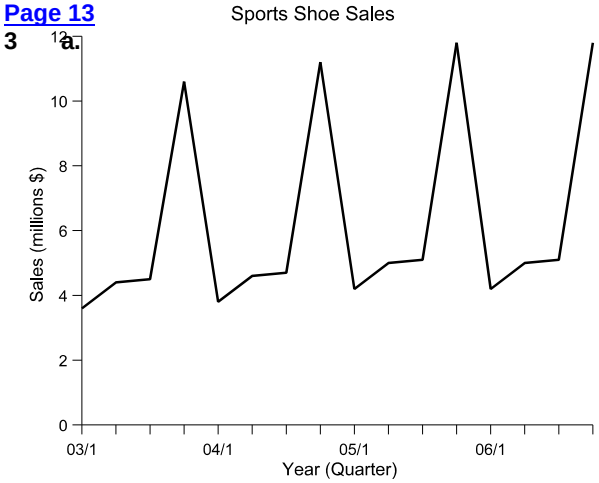
99/00	00/01	01/02	02/03	03/04	04/05	05/06
1056	1103	1182	1230	1251	1252	1201
MMean	1114	1172	1221	1244	1235	



b. The four components of time series data are trend, cyclical, seasonal and irregular. The trend component reflects the long term movement of the data. In the data given, there is an upward trend (probably caused by increasing population increased consumption and productivity etc. The cyclical component refers to regular cycles that occur. In the data given, there is a steady increase until 2005/2006 when there is a downturn. If the data were continued you would probably find that the downturn would continue for a while and then recover. This is known in economic circles as the boom - recession - recovery - boom cycle. However more data is needed to confirm it. The seasonal component refers to regular cycles in the time series occurring at the same time each year. (e.g., increased retail sales in Nov & Dec.) As the series of data given is for annual figures, we can not recover from them any seasonal fluctuations.
Irregular fluctuations occur because of unpredictable factors. There is no evidence of irregular functions in this graph.

Page 13

3 a.



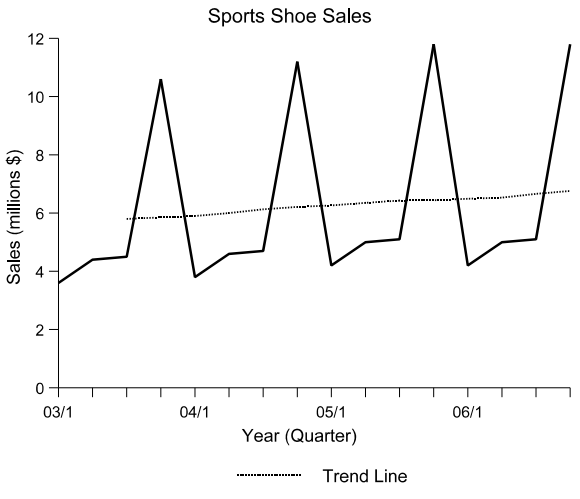
b. You could not just draw a line of best fit through the graph as the fourth quarter of each year would cause an overestimate for the other three quarters. However you could consider drawing trend lines through each of the seasonal values to predict future sales for individual quarters.

Page 14

c.

Quarter		Sales	Yearly Total	Moving Mean	Centred Mean	Variation
2003	1	3.6				
	2	4.4				
			23.1	5.78		
	3	4.5			5.8	-1.30
			23.3	5.83		
	4	10.6			5.85	4.75
			23.5	5.88		
2004	1	3.8			5.9	-2.10
			23.7	5.93		
	2	4.6			6	-1.40
			24.3	6.08		
	3	4.7			6.13	-1.43
			24.7	6.18		
	4	11.2			6.21	4.99
			25.0	6.25		
2005	1	4.2			6.26	-2.06
			25.1	6.28		
	2	4.9			6.35	-1.45
			25.7	6.43		
	3	4.8			6.43	-1.63
			25.7	6.43		
	4	11.8			6.44	5.36
			25.8	6.45		
2006	1	4.2			6.49	-2.29
			26.1	6.53		
	2	5.0			6.53	-1.53
			26.1	6.53		
	3	5.1				
	4	11.8				

The graph below has the moving averages added



Page 14 cont

Table of Variations

	Q1	Q2	Q3	Q4
2003			-1.30	4.75
2004	-2.10	-1.40	-1.43	4.99
2005	-2.06	-1.45	-1.63	5.36
2006	-2.29	-1.53		
Total	-6.45	-4.38	-4.36	15.10
Average	-2.15	-1.46	-1.45	5.03

To calculate the third and fourth quarters deseasonalised sales, subtract the seasonal factor: The calculated figures allow you to extend the trend line.

5.1 - (-1.45) = 6.55 (6.6 1 dp)

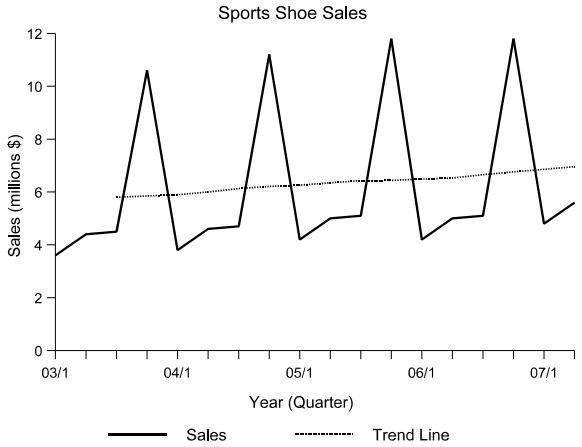
and 11.8 -(5.03) = 6.77 (6.8 1 dp)

Sales of shoes with the trend line projected and the estimated sales for 2007 Quarter 1 & Quarter 2 is below.

Reading from the projected trend line, you should get the values of Q1, 2007 = 6.9 and Q2, 2007 = 7.1 (each in $millions). Note: to get these figures your graph will need to be larger and have a better scale than the one found in these answers which is for informational purposes only.

These figures are then "seasonalised" to be:
6.9 + (- 2.15) = 4.8 and 7.1 + (- 1.46) = 5.6 (1 dp).

These figures are now added to the original graph and it is roughly representative of what has gone before.



Page 15

5. **a.** (3054 + 3455 + 2869) ÷ 3 = 3126
3455 - 3126 = 329

b. July: = (266 + 329 + 311) ÷ 3
= 302
Nov = ((-116) + (-138) + (-164) + (-143)) ÷ 4
= -140.25
= -140
March = ((-132) + (-160) + (-159) + (-171)) ÷ 4
= -155.5
= -156

c. 2879 - 302 = 2577
The unemployment numbers are actually declining. This is evident by the decline of the moving mean (i.e. the trend line).

d. March 31, 2007 = 2658 + 3(-65) + (-156)
= 2307
July 31, 2007 = 2658 + 4(-65) + (302)
= 2700

Page 16

6. **a.** 425 - 445.875 = -20.88

b. (15 + 8.25) ÷ 2 = 11.63

c. December 2004 - December 2005 (5 quarters)
i.e (461+387+451+425)/4 = 431
(387+451+425+501)/4 = 441
(431+441)/2 = 436

d. **i.** 478 + (-36.38) = 441.62
= $441 620

ii. Between December 2003 and March 2008 there are 17 quarters. To predict the actual values you must also add the seasonal effect.
S = 389 + 6.5Q + seasonal effect
= 389 + (6.5 × 17) + (-36.38)
= 463.12
= $463 120

Page 17

7. **a.** Each of the peaks and troughs are found at regular intervals.

b. **i.** n = 4

ii. The centred moving mean enables the values to correspond to each quarter value rather than be positioned midway between.

iii. Mar (07) t = 21
T = 182.6 + (1.57 × 21) + (-16.88)
= 198.69
= $198 690
Jun(07) t = 22
T = 182.6 + (1.57 × 22) + (-31.44)
= 185.70
= $185 700
Sept (07) t = 23
T = 182.6 + (1.57 × 23) + 12.97
= 231.68
= $231 680
Dec (07) t = 24
T = 182.6 + (1.57 × 24) + 37.34
= 257.62
= $257 620

Page 18

8. **a.** **i.** -229 **ii.** 6650 - 58 = 6592

b. The centred moving mean as it deseasonalises the data and shows whether it is trending upwards or downwards.

c. The June seasonally adjusted value of 7576 is below the centred moving mean of 7619 therefore the stock value was lower than expected.
OR
You could also say the that the predicted value was 7619 + (-26) = 7593. The actual value is still lower than the predicted value.

d. March 06, June 06, Sept 06, Dec 06 = 4 periods
= 8004 + (4 × 205) + 58
= 8882

Page 19

9. **a.** Moving Mean = (0.56 + 0.483 + 0.707 + 0.676) ÷ 4
= 0.6065
Centred Moving Mean = (0.7690 + 0.7610) ÷ 2
= 0.7650
Individual Seasonal Effects = 0.883 - 0.7474
= 0.1356

b. Summer seasonal effect
= (0.11 + 0.1375 + 0.0681) ÷ 3
= 0.1052

c. 0.707 - 0.1052 = 0.6018
This is below the centred moving mean trend line (0.6389) meaning that occupancy was lower then expected OR you could also add the summer seasonal effect to the centred moving mean 0.1052 + 0.6380 = 0.7434. This is the predicted value which is higher than the actual value of 0.707 which confirms occupancy was lower than expected.

Page 20

d. The centred moving mean gives the long term trend which is a decreasing rate of occupancy.

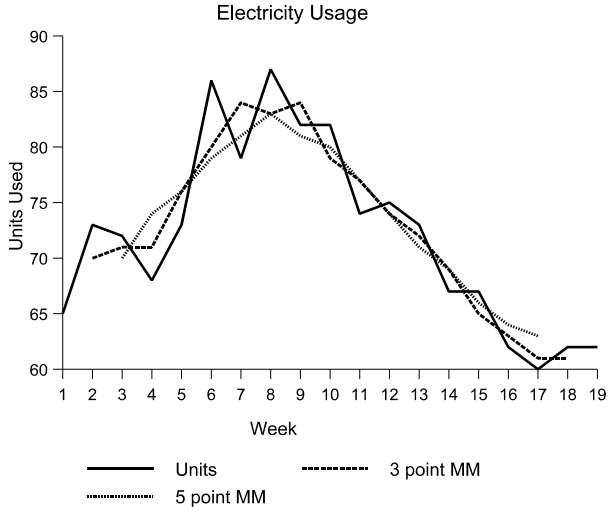
e. Look at whether the seasonally adjusted rates are above or below the moving mean trend line. Summer and Spring are below therefore occupancy was worse then expected, Autumn and Winter are above therefore must have had better occupancy than expected.

f. The Summer Seasonal Effect = 0.1052
Using the moving mean for Spring 2005
(0.5993) + (5 × -0.01) + (0.1052) = 0.6545

Page 21

10.

Week	Units Used	Moving Means 3 point	5 point
1	65		
2	73	70.00	
3	72	71.00	70.20
4	68	71.00	74.40
5	73	75.67	75.60
6	86	79.33	78.60
7	79	84.00	81.40
8	87	82.67	83.20
9	82	83.67	80.80
10	82	79.33	80.00
11	74	77.00	77.20
12	75	74.00	74.20
13	73	71.67	71.20
14	67	69.00	68.80
15	67	65.33	65.80
16	62	63.00	63.60
17	60	61.33	62.60
18	62	61.33	
19	62		



The plots are very similar however the 5 point moving mean smooths the data less drastically than the 3 point. A moving mean gives a smooth trend line version of the time series. This graph does have a trend upwards and then downwards indicating a possible start of winter then moving onto summer. However you really need at least two lots of 52 weeks of data to identify any real seasonal differences.

Page 24

Year	Value (Y)	Dev (X)	XY	X^2
2000	10	1	10	1
2001	15	2	30	4
2002	21	3	63	9
2003	25	4	100	16
2004	32	5	160	25
2005	40	6	240	36
6	**143**	**21**	**603**	**91**

The normal equations are:

$\sum Y = Na + b\sum X$

and $\sum XY = a\sum X + b\sum X^2$

Substituting values: 143 = 6a + 21b
603 = 21a + 91b
Multiplying equation 1 by 7: 1001 = 42a + 147b
Multiplying equation 1 by 2: 1206 = 42a + 182b
Subtracting -205 = - 35b
b = 5.857 (3 DP)

Substituting b into equation 1
143 = 6a + (21×5.857)
143 = 6a + 122.997
a = (143 - 122.997) ÷ 6
a = 3.3 (1 DP)
∴ equation of trend line is Y = 3.3 + 5.857X
∴ trend values are:

	2000	2001	2002	2003	2004	2005
Actual	10	15	21	25	32	40
Trend	9.19	15.05	20.90	26.76	32.62	38.48

Page 25

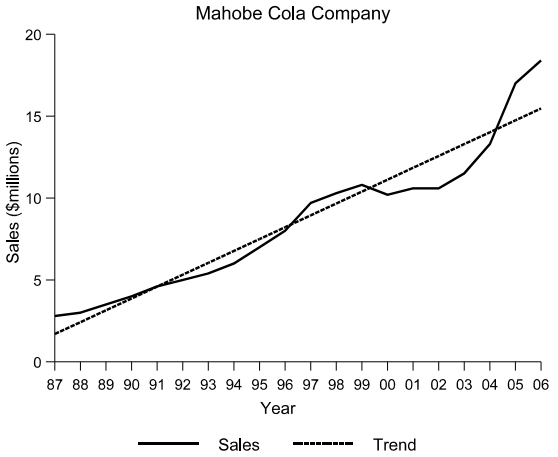
2.

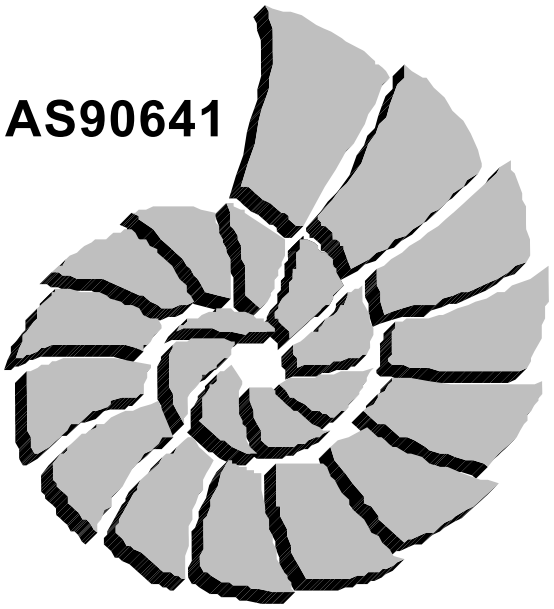
Year	Value (Y)	Deviation (X)	XY	X^2
1987	2.8	0	0.0	0
1988	3.0	1	3.0	1
1989	3.5	2	7.0	4
1990	4.0	3	12.0	9
1991	4.6	4	18.4	16
1992	5.0	5	25.0	25
1993	5.4	6	32.4	36
1994	6.0	7	42.0	49
1995	7.0	8	56.0	64
1996	8.0	9	72.0	81
1997	9.7	10	97.0	100
1998	10.3	11	113.3	121
1999	10.8	12	129.6	144
2000	10.2	13	132.6	169
2001	10.6	14	148.4	196
2002	10.6	15	159.0	225
2003	11.5	16	184.0	256
2004	13.3	17	226.1	289
2005	17.0	18	306.0	324
2006	18.4	19	349.6	361
∑ = 20	171.7	190	2113.4	2470

Substituting the values into the normal equations:
171.7 = 20a + 190b
2113.4 = 190a + 2470b
To calculate These figures can now be put into the Equation Function of a fx9750G PLUS calculator

	a	b	c
1	20	190	171.7
2	190	2470	2113.4

The solved values are: (a) Intercept = 1.696
(b) Gradient = 0.725
∴ the trend line is Y = 1.696 + 0.725X
The slope of the trend line indicates that net sales are increasing at a rate of $0.725 (million) each year. However looking at the graph you will note that in recent years a marked difference has occurred in the sales. A quadratic or exponential trend line could present this model better.





Exemplar Material (1) For

Time Series

Kim Freeman

This book provides an exemplar for
NZQA, Level 3 Mathematics
Statistics and Modelling 3.1
Determine the Trend for Time Series

Exemplar Material (1) For Time Series
Published 2007 by:
Mahobe Resources (NZ) Ltd
P.O. Box 109-760
Newmarket, Auckland 1149
New Zealand

admin@mahobe.co.nz
www.mahobe.co.nz

ISBN 0-9583564-8-3
ISBN13 9780958356480

'Exemplar Material (1) For Time Series' provides material for the New Zealand Qualifications Authority, NCEA Level 3 Mathematics Achievement Standard 3.1 of Statistics and Modelling. The standard is AS90641: Determine the trend for time series data.

Mahobe Resources are grateful for the students who gave permission to use their work as resources. The marking and grading of these was the responsibility of the teachers of these students and has not been considered by the publisher. Any alterations to the script was only to help clarify areas in which the publisher believed may assist the reader.

Mahobe Resources (NZ) Ltd welcomes feedback on the book and its usefulness.

Time Series
Internal Assessment Resource

The data below gives monthly data (in 100s) of the numbers of containers being processed through Port Mahobe(NZ) Ltd each month from 1997 through to 2006.

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Jan	262	259	271	251	298	260	275	315	354	417
Feb	295	276	241	231	283	291	321	342	365	408
Mar	333	310	301	252	315	307	352	370	389	416
Apr	252	238	265	293	287	293	322	316	198	398
May	274	270	255	278	301	279	309	361	366	397
Jun	245	292	301	447	185	287	314	320	389	452
Jul	377	289	278	216	368	344	299	324	341	423
Aug	291	289	262	247	310	359	355	320	413	456
Sep	273	273	246	267	313	250	324	344	387	356
Oct	266	271	249	281	312	368	310	300	384	479
Nov	286	272	246	297	325	359	339	350	415	425
Dec	285	284	221	288	326	345	320	333	328	499

Your task is to:

1. For achievement: Determine the trend for the time series
   a. Develop a table containing the raw data, an appropriate trend line and the individual monthly effects.
   b. Draw a graph of the raw data and your trend line on the same axes.
   c. Give a quantitative (numerical) description of the trend and give reasons that might account for this trend.

2. For merit: Analyse time series data to make a forecast.
   a. Make a prediction for the first 6 months of 2007.
   b. You must show how you established these forecasts based on estimates of the trend for the smoothed data and estimates of monthly effects.

3. For excellence: Report on the validity of the analysis.
   a. Write a report on your analysis. Your report must include justified comments on ALL of the following:
      i. Relevance and usefulness of the forecast.
      ii. Features of the time series data
      iii. Appropriateness of the model
      iv. Improvements to the model.
      v. Limitations of the analysis.

A table with a 5 point centred moving mean and individual monthly variations is below.

Table 1 Year	Month	Value	5 Point Moving Mean	Variation		Year / Month		Value	Moving Mean	Variation
1997	Jan	262				2002	Jan	260	301.8	-41.8
	Feb	295					Feb	291	295.4	-4.4
	Mar	333	283.2	49.8			Mar	307	286	21
	Apr	252	279.8	-27.8			Apr	293	291.4	1.6
	May	274	296.2	-22.2			May	279	302	-23
	Jun	245	287.8	-42.8			Jun	287	312.4	-25.4
	Jul	377	292	85			Jul	344	303.8	40.2
	Aug	291	290.4	0.6			Aug	359	321.6	37.4
	Sep	273	298.6	-25.6			Sep	250	336	-86
	Oct	266	280.2	-14.2			Oct	368	336.2	31.8
	Nov	286	273.8	12.2			Nov	359	319.4	39.6
	Dec	285	274.4	10.6			Dec	345	333.6	11.4
1998	Jan	259	283.2	-24.2		2003	Jan	275	330.4	-55.4
	Feb	276	273.6	2.4			Feb	321	323	-2
	Mar	310	270.6	39.4			Mar	352	315.8	36.2
	Apr	238	277.2	-39.2			Apr	322	323.6	-1.6
	May	270	279.8	-9.8			May	309	319.2	-10.2
	Jun	292	275.6	16.4			Jun	314	319.8	-5.8
	Jul	289	282.6	6.4			Jul	299	320.2	-21.2
	Aug	289	282.8	6.2			Aug	355	320.4	34.6
	Sep	273	278.8	-5.8			Sep	324	325.4	-1.4
	Oct	271	277.8	-6.8			Oct	310	329.6	-19.6
	Nov	272	274.2	-2.2			Nov	339	321.6	17.4
	Dec	284	267.8	16.2			Dec	320	325.2	-5.2
1999	Jan	271	273.8	-2.8		2004	Jan	315	337.2	-22.2
	Feb	241	272.4	-31.4			Feb	342	332.6	9.4
	Mar	301	266.6	34.4			Mar	370	340.8	29.2
	Apr	265	272.6	-7.6			Apr	316	341.8	-25.8
	May	255	280	-25			May	361	338.2	22.8
	Jun	301	272.2	28.8			Jun	320	328.2	-8.2
	Jul	278	268.4	9.6			Jul	324	333.8	-9.8
	Aug	262	267.2	-5.2			Aug	320	321.6	-1.6
	Sep	246	256.2	-10.2			Sep	344	327.6	16.4
	Oct	249	244.8	4.2			Oct	300	329.4	-29.4
	Nov	246	242.6	3.4			Nov	350	336.2	13.8
	Dec	221	239.6	-18.6			Dec	333	340.4	-7.4
2000	Jan	251	240.2	10.8		2005	Jan	354	358.2	-4.2
	Feb	231	249.6	-18.6			Feb	365	327.8	37.2
	Mar	252	261	-9			Mar	389	334.4	54.6
	Apr	293	300.2	-7.2			Apr	198	341.4	-143.4
	May	278	297.2	-19.2			May	366	336.6	29.4
	Jun	447	296.2	150.8			Jun	389	341.4	47.6
	Jul	216	291	-75			Jul	341	379.2	-38.2
	Aug	247	291.6	-44.6			Aug	413	382.8	30.2
	Sep	267	261.6	5.4			Sep	387	388	-1
	Oct	281	276	5			Oct	384	385.4	-1.4
	Nov	297	286.2	10.8			Nov	415	386.2	28.8
	Dec	288	289.4	-1.4			Dec	328	390.4	-62.4
2001	Jan	298	296.2	1.8		2006	Jan	417	396.8	20.2
	Feb	283	294.2	-11.2			Feb	408	393.4	14.6
	Mar	315	296.8	18.2			Mar	416	407.2	8.8
	Apr	287	274.2	12.8			Apr	398	414.2	-16.2
	May	301	291.2	9.8			May	397	417.2	-20.2
	Jun	185	290.2	-105.2			Jun	452	425.2	26.8
	Jul	368	295.4	72.6			Jul	423	416.8	6.2
	Aug	310	297.6	12.4			Aug	456	433.2	22.8
	Sep	313	325.6	-12.6			Sep	356	427.8	-71.8
	Oct	312	317.2	-5.2			Oct	479	443	36
	Nov	325	307.2	17.8			Nov	425		
	Dec	326	302.8	23.2			Dec	499		

A table with a 7 point centred moving mean and individual monthly variations is below.

Table 2 Year	Month	Value	7 Point Moving Mean	Variation		Year / Month		Value	Moving Mean	Variation
1997	Jan	262				2002	Jan	260	302	-42
	Feb	295					Feb	291	297.29	-6.29
	Mar	333					Mar	307	291.86	15.14
	Apr	252	291.14	-39.14			Apr	293	294.43	-1.43
	May	274	295.29	-21.29			May	279	308.57	-29.57
	Jun	245	292.14	-47.14			Jun	287	302.71	-15.71
	Jul	377	282.57	94.43			Jul	344	311.43	32.57
	Aug	291	287.43	3.57			Aug	359	320.86	38.14
	Sep	273	289	-16			Sep	250	330.29	-80.29
	Oct	266	291	-25			Oct	368	328.57	39.43
	Nov	286	276.57	9.43			Nov	359	325.29	33.71
	Dec	285	279.29	5.71			Dec	345	324.29	20.71
1998	Jan	259	274.29	-15.29		2003	Jan	275	334.57	-59.57
	Feb	276	274.86	1.14			Feb	321	326.14	-5.14
	Mar	310	275.71	34.29			Mar	352	319.71	32.29
	Apr	238	276.29	-38.29			Apr	322	313.14	8.86
	May	270	280.57	-10.57			May	309	324.57	-15.57
	Jun	292	280.14	11.86			Jun	314	325	-11
	Jul	289	274.57	14.43			Jul	299	319	-20
	Aug	289	279.43	9.57			Aug	355	321.43	33.57
	Sep	273	281.43	-8.43			Sep	324	323	1
	Oct	271	278.43	-7.43			Oct	310	323.14	-13.14
	Nov	272	271.57	0.43			Nov	339	329.29	9.71
	Dec	284	273.29	10.71			Dec	320	331.43	-11.43
1999	Jan	271	272.14	-1.14		2004	Jan	315	330.29	-15.29
	Feb	241	269.86	-28.86			Feb	342	337.57	4.43
	Mar	301	274	27			Mar	370	334.86	35.14
	Apr	265	273.14	-8.14			Apr	316	335.43	-19.43
	May	255	271.86	-16.86			May	361	336.14	24.86
	Jun	301	272.57	28.43			Jun	320	336.43	-16.43
	Jul	278	265.14	12.86			Jul	324	326.43	-2.43
	Aug	262	262.43	-0.43			Aug	320	331.29	-11.29
	Sep	246	257.57	-11.57			Sep	344	327.29	16.71
	Oct	249	250.43	-1.43			Oct	300	332.14	-32.14
	Nov	246	243.71	2.29			Nov	350	338	12
	Dec	221	242.29	-21.29			Dec	333	347.86	-14.86
2000	Jan	251	249	2		2005	Jan	354	327	27
	Feb	231	253.14	-22.14			Feb	365	336.43	28.57
	Mar	252	281.86	-29.86			Mar	389	342	47
	Apr	293	281.14	11.86			Apr	198	343.14	-145.14
	May	278	280.57	-2.57			May	366	351.57	14.43
	Jun	447	285.71	161.29			Jun	389	354.71	34.29
	Jul	216	289.86	-73.86			Jul	341	354	-13
	Aug	247	290.43	-43.43			Aug	413	385	28
	Sep	267	291.86	-24.86			Sep	387	379.57	7.43
	Oct	281	270.57	10.43			Oct	384	383.57	0.43
	Nov	297	280.14	16.86			Nov	415	393.14	21.86
	Dec	288	289.86	-1.86			Dec	328	393.57	-65.57
2001	Jan	298	292.71	5.29		2006	Jan	417	395.14	21.86
	Feb	283	295.57	-12.57			Feb	408	397	11
	Mar	315	279.57	35.43			Mar	416	402.29	13.71
	Apr	287	291	-4			Apr	398	415.86	-17.86
	May	301	292.71	8.29			May	397	421.43	-24.43
	Jun	185	297	-112			Jun	452	414	38
	Jul	368	296.57	71.43			Jul	423	423	0
	Aug	310	302	8			Aug	456	426.86	29.14
	Sep	313	305.57	7.43			Sep	356	441.43	-85.43
	Oct	312	316.29	-4.29			Oct	479		
	Nov	325	305.29	19.71			Nov	425		
	Dec	326	304.86	21.14			Dec	499		

Figure 1

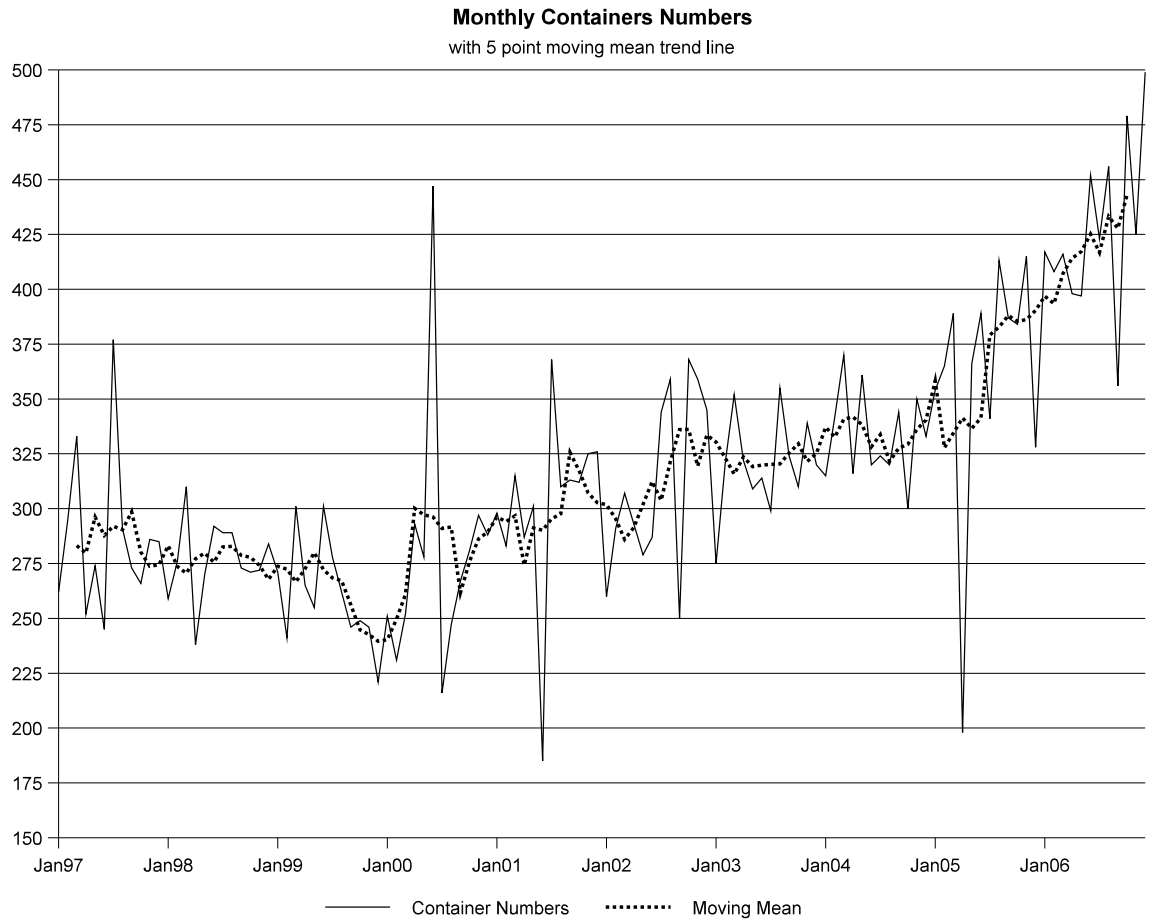


Figure 2

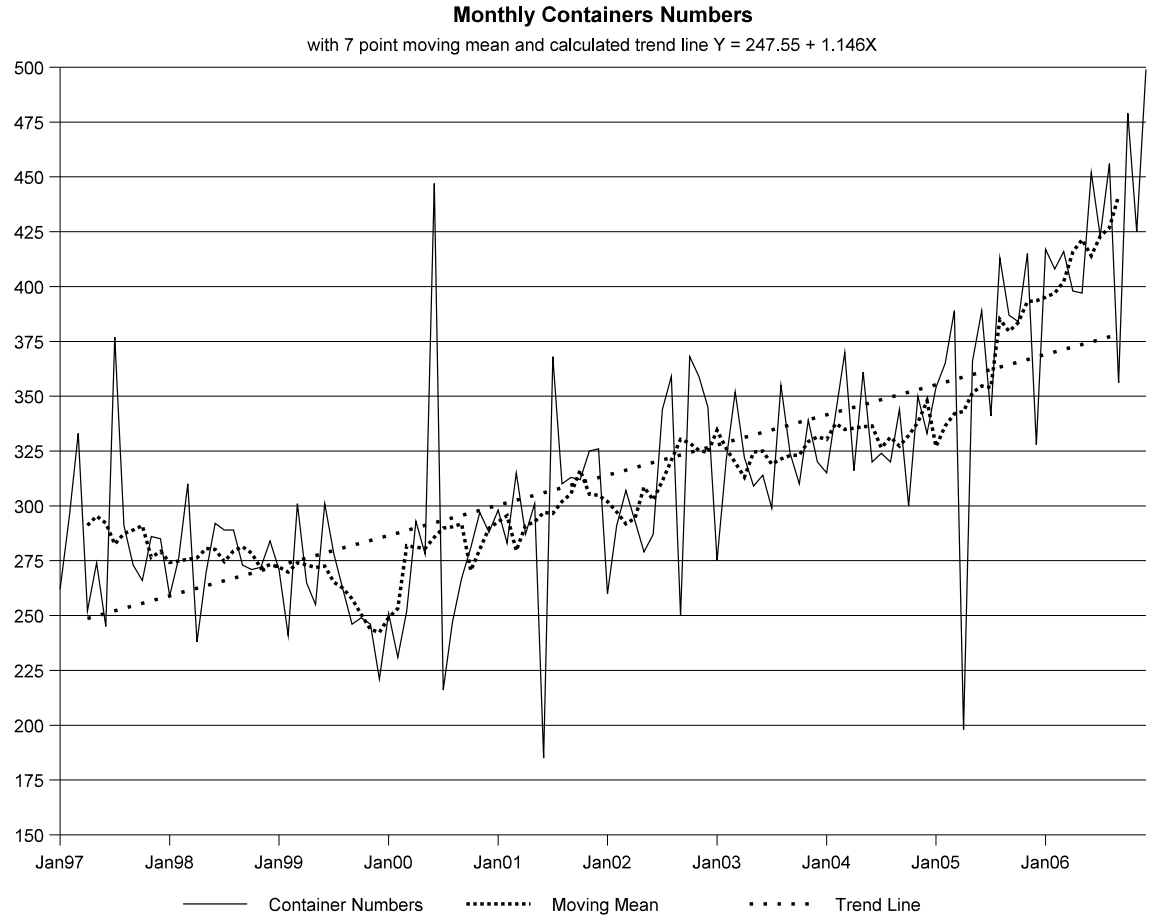


Table 3 (below) gives the data and least square calculations.

Table 3 Year / Month		Value (Y)	Deviation from 1997 (X)	XY	X^2		Year / Month		Value	Deviation from 1997 (X)	XY	X^2
1997	Jan	262	0	0	0		2002	Jan	260	60	15600	3600
	Feb	295	1	295	1			Feb	291	61	17751	3721
	Mar	333	2	666	4			Mar	307	62	19034	3844
	Apr	252	3	756	9			Apr	293	63	18459	3969
	May	274	4	1096	16			May	279	64	17856	4096
	Jun	245	5	1225	25			Jun	287	65	18655	4225
	Jul	377	6	2262	36			Jul	344	66	22704	4356
	Aug	291	7	2037	49			Aug	359	67	24053	4489
	Sep	273	8	2184	64			Sep	250	68	17000	4624
	Oct	266	9	2394	81			Oct	368	69	25392	4761
	Nov	286	10	2860	100			Nov	359	70	25130	4900
	Dec	285	11	3135	121			Dec	345	71	24495	5041
1998	Jan	259	12	3108	144		2003	Jan	275	72	19800	5184
	Feb	276	13	3588	169			Feb	321	73	23433	5329
	Mar	310	14	4340	196			Mar	352	74	26048	5476
	Apr	238	15	3570	225			Apr	322	75	24150	5625
	May	270	16	4320	256			May	309	76	23484	5776
	Jun	292	17	4964	289			Jun	314	77	24178	5929
	Jul	289	18	5202	324			Jul	299	78	23322	6084
	Aug	289	19	5491	361			Aug	355	79	28045	6241
	Sep	273	20	5460	400			Sep	324	80	25920	6400
	Oct	271	21	5691	441			Oct	310	81	25110	6561
	Nov	272	22	5984	484			Nov	339	82	27798	6724
	Dec	284	23	6532	529			Dec	320	83	26560	6889
1999	Jan	271	24	6504	576		2004	Jan	315	84	26460	7056
	Feb	241	25	6025	625			Feb	342	85	29070	7225
	Mar	301	26	7826	676			Mar	370	86	31820	7396
	Apr	265	27	7155	729			Apr	316	87	27492	7569
	May	255	28	7140	784			May	361	88	31768	7744
	Jun	301	29	8729	841			Jun	320	89	28480	7921
	Jul	278	30	8340	900			Jul	324	90	29160	8100
	Aug	262	31	8122	961			Aug	320	91	29120	8281
	Sep	246	32	7872	1024			Sep	344	92	31648	8464
	Oct	249	33	8217	1089			Oct	300	93	27900	8649
	Nov	246	34	8364	1156			Nov	350	94	32900	8836
	Dec	221	35	7735	1225			Dec	333	95	31635	9025
2000	Jan	251	36	9036	1296		2005	Jan	354	96	33984	9216
	Feb	231	37	8547	1369			Feb	365	97	35405	9409
	Mar	252	38	9576	1444			Mar	389	98	38122	9604
	Apr	293	39	11427	1521			Apr	198	99	19602	9801
	May	278	40	11120	1600			May	366	100	36600	10000
	Jun	447	41	18327	1681			Jun	389	101	39289	10201
	Jul	216	42	9072	1764			Jul	341	102	34782	10404
	Aug	247	43	10621	1849			Aug	413	103	42539	10609
	Sep	267	44	11748	1936			Sep	387	104	40248	10816
	Oct	281	45	12645	2025			Oct	384	105	40320	11025
	Nov	297	46	13662	2116			Nov	415	106	43990	11236
	Dec	288	47	13536	2209			Dec	328	107	35096	11449
2001	Jan	298	48	14304	2304		2006	Jan	417	108	45036	11664
	Feb	283	49	13867	2401			Feb	408	109	44472	11881
	Mar	315	50	15750	2500			Mar	416	110	45760	12100
	Apr	287	51	14637	2601			Apr	398	111	44178	12321
	May	301	52	15652	2704			May	397	112	44464	12544
	Jun	185	53	9805	2809			Jun	452	113	51076	12769
	Jul	368	54	19872	2916			Jul	423	114	48222	12996
	Aug	310	55	17050	3025			Aug	456	115	52440	13225
	Sep	313	56	17528	3136			Sep	356	116	41296	13456
	Oct	312	57	17784	3249			Oct	479	117	56043	13689
	Nov	325	58	18850	3364			Nov	425	118	50150	13924
	Dec	326	59	19234	3481			Dec	499	119	59381	14161
							sums	120	37901	7140	2426764	568820
							Intercept = 244.91			Gradient = 1.1921		

Table 4 (below) uses the trend line from Excel least square calculations: Y = 244.91 + 1.1921X . It gives the calculated trend value and also gives the monthly variations from March 1997 to December 2005. The data was taken from the Table of Least Squares.

Table 4 Year / Month		Value (Y)	Deviation from 1997 (X)	Forecast	Deviation		Year / Month		Value	Deviation from 1997 (X)	Forecast	Deviation
1997	Jan	262	0	244.91	17.09		2002	Jan	260	60	316.44	-56.44
	Feb	295	1	246.1	48.9			Feb	291	61	317.63	-26.63
	Mar	333	2	247.29	85.71			Mar	307	62	318.82	-11.82
	Apr	252	3	248.49	3.51			Apr	293	63	320.01	-27.01
	May	274	4	249.68	24.32			May	279	64	321.2	-42.2
	Jun	245	5	250.87	-5.87			Jun	287	65	322.4	-35.4
	Jul	377	6	252.06	124.94			Jul	344	66	323.59	20.41
	Aug	291	7	253.25	37.75			Aug	359	67	324.78	34.22
	Sep	273	8	254.45	18.55			Sep	250	68	325.97	-75.97
	Oct	266	9	255.64	10.36			Oct	368	69	327.16	40.84
	Nov	286	10	256.83	29.17			Nov	359	70	328.36	30.64
	Dec	285	11	258.02	26.98			Dec	345	71	329.55	15.45
1998	Jan	259	12	259.22	-0.22		2003	Jan	275	72	330.74	-55.74
	Feb	276	13	260.41	15.59			Feb	321	73	331.93	-10.93
	Mar	310	14	261.6	48.4			Mar	352	74	333.13	18.87
	Apr	238	15	262.79	-24.79			Apr	322	75	334.32	-12.32
	May	270	16	263.98	6.02			May	309	76	335.51	-26.51
	Jun	292	17	265.18	26.82			Jun	314	77	336.7	-22.7
	Jul	289	18	266.37	22.63			Jul	299	78	337.89	-38.89
	Aug	289	19	267.56	21.44			Aug	355	79	339.09	15.91
	Sep	273	20	268.75	4.25			Sep	324	80	340.28	-16.28
	Oct	271	21	269.94	1.06			Oct	310	81	341.47	-31.47
	Nov	272	22	271.14	0.86			Nov	339	82	342.66	-3.66
	Dec	284	23	272.33	11.67			Dec	320	83	343.85	-23.85
1999	Jan	271	24	273.52	-2.52		2004	Jan	315	84	345.05	-30.05
	Feb	241	25	274.71	-33.71			Feb	342	85	346.24	-4.24
	Mar	301	26	275.9	25.1			Mar	370	86	347.43	22.57
	Apr	265	27	277.1	-12.1			Apr	316	87	348.62	-32.62
	May	255	28	278.29	-23.29			May	361	88	349.81	11.19
	Jun	301	29	279.48	21.52			Jun	320	89	351.01	-31.01
	Jul	278	30	280.67	-2.67			Jul	324	90	352.2	-28.2
	Aug	262	31	281.87	-19.87			Aug	320	91	353.39	-33.39
	Sep	246	32	283.06	-37.06			Sep	344	92	354.58	-10.58
	Oct	249	33	284.25	-35.25			Oct	300	93	355.78	-55.78
	Nov	246	34	285.44	-39.44			Nov	350	94	356.97	-6.97
	Dec	221	35	286.63	-65.63			Dec	333	95	358.16	-25.16
2000	Jan	251	36	287.83	-36.83		2005	Jan	354	96	359.35	-5.35
	Feb	231	37	289.02	-58.02			Feb	365	97	360.54	4.46
	Mar	252	38	290.21	-38.21			Mar	389	98	361.74	27.26
	Apr	293	39	291.4	1.6			Apr	198	99	362.93	-164.93
	May	278	40	292.59	-14.59			May	366	100	364.12	1.88
	Jun	447	41	293.79	153.21			Jun	389	101	365.31	23.69
	Jul	216	42	294.98	-78.98			Jul	341	102	366.5	-25.5
	Aug	247	43	296.17	-49.17			Aug	413	103	367.7	45.3
	Sep	267	44	297.36	-30.36			Sep	387	104	368.89	18.11
	Oct	281	45	298.55	-17.55			Oct	384	105	370.08	13.92
	Nov	297	46	299.75	-2.75			Nov	415	106	371.27	43.73
	Dec	288	47	300.94	-12.94			Dec	328	107	372.46	-44.46
2001	Jan	298	48	302.13	-4.13		2006	Jan	417	108	373.66	43.34
	Feb	283	49	303.32	-20.32			Feb	408	109	374.85	33.15
	Mar	315	50	304.52	10.48			Mar	416	110	376.04	39.96
	Apr	287	51	305.71	-18.71			Apr	398	111	377.23	20.77
	May	301	52	306.9	-5.9			May	397	112	378.43	18.57
	Jun	185	53	308.09	-123.09			Jun	452	113	379.62	72.38
	Jul	368	54	309.28	58.72			Jul	423	114	380.81	42.19
	Aug	310	55	310.48	-0.48			Aug	456	115	382	74
	Sep	313	56	311.67	1.33			Sep	356	116	383.19	-27.19
	Oct	312	57	312.86	-0.86			Oct	479	117	384.39	94.61
	Nov	325	58	314.05	10.95			Nov	425	118	385.58	39.42
	Dec	326	59	315.24	10.76			Dec	499	119	386.77	112.23

Figure 3 (below) gives a graph of the data with the calculated trend line: Y = 244.91 + 1.1921X .

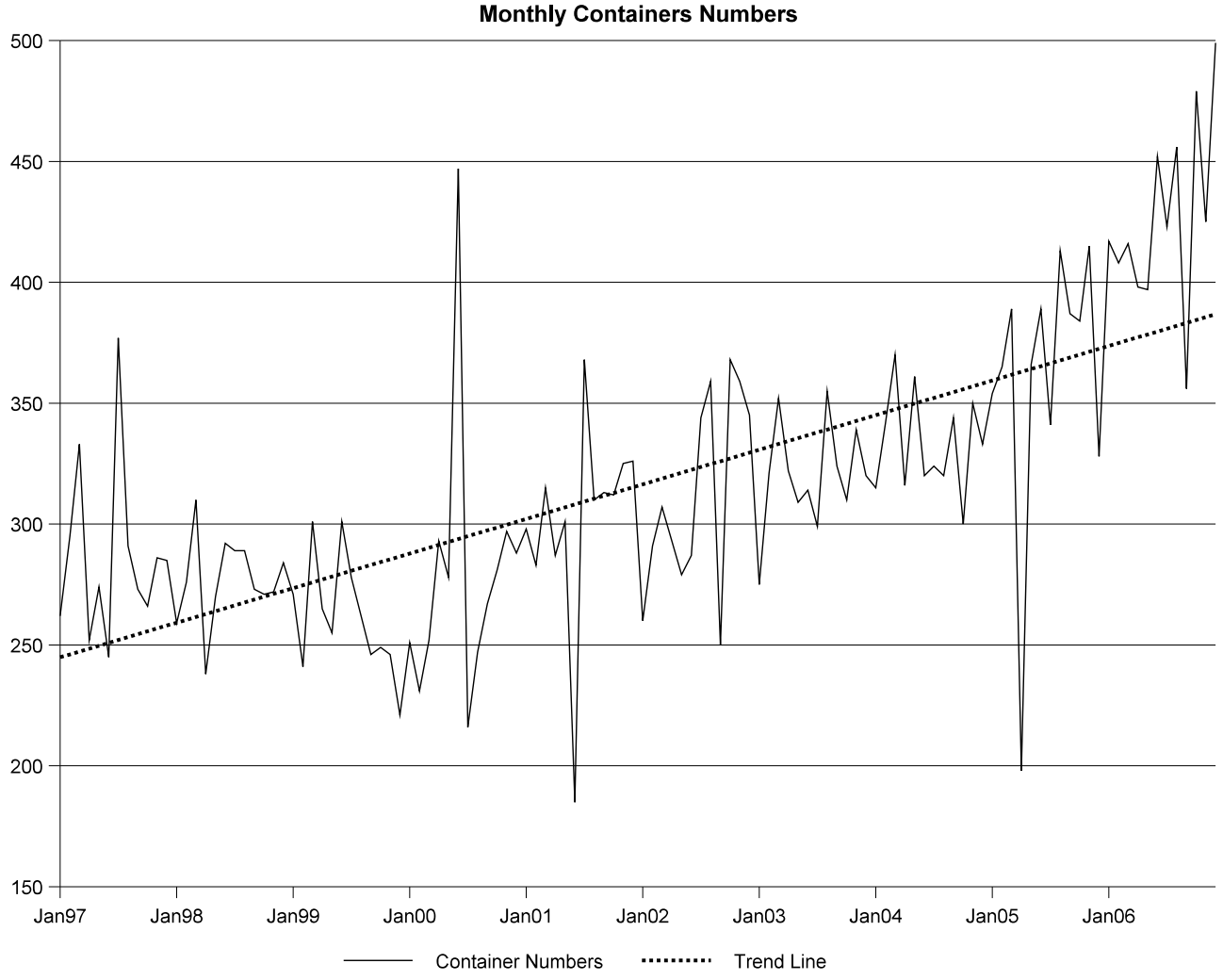


Table 5 (below) calculates the monthly variations from March 1997 to December 2005. The data is taken from the Table of 7 point Moving Means.

Table 5	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Average
Jan		-15.29	-1.1	2	5.3	-42	-59.6	-15.3	27	21.9	-8.57
Feb		1.1	-28.9	-22.1	-12.6	-6.3	-5.1	4.4	28.6	11	-3.32
Mar		34.3	27	-29.9	35.4	15.1	32.29	35.1	47	13.7	23.33
Apr	-39.1	-38.3	-8.14	11.9	-4	-1.4	8.9	-19.4	-145.14	-17.9	-25.26
May	-21.3	-10.6	-16.9	-2.6	8.3	-29.6	-15.6	24.9	14.4	-24.4	-7.34
Jun	-47.1	11.9	28.4	161.3	-112	-15.7	-11	-16.4	34.3	38	7.17
Jul	94.4	14.4	12.9	-73.9	71.4	32.6	-20	-2.4	-13	0	11.64
Aug	3.57	9.6	-0.4	-43.4	8	38.1	33.6	-11.3	28	29.1	9.49
Sep	-16	-8.4	-11.6	-24.9	7.4	-80.3	1	16.7	7.4	-85.4	-19.41
Oct	-25	-7.4	-1.4	10.4	-4.3	39.4	-13.1	-32.1	0.4		-3.68
Nov	9.43	0.43	2.3	16.9	19.7	33.7	9.7	12	21.9		14.01
Dec	5.71	10.7	-21.3	-1.86	21.1	20.7	-11.4	-14.9	-65.6		-6.31

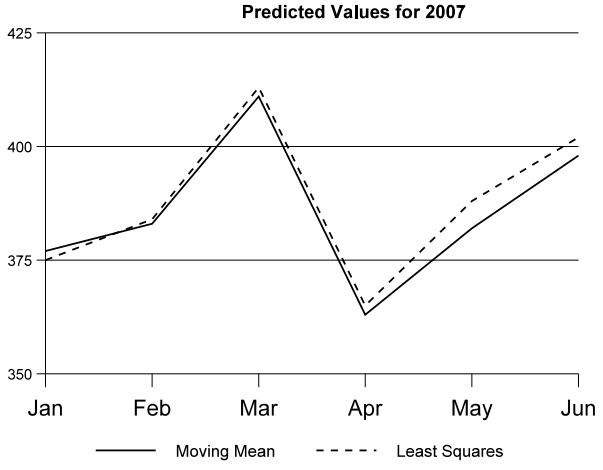
Table 6 (below) gives the average monthly variations from March 1997 to December 2005. The data is taken from the Table of Least Squares.

Table 6	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	Average
Jan	17.1	-0.2	-2.5	-36.8	-4.1	-56.4	-55.7	-30.0	-5.4	43.3	-13.07
Feb	48.9	15.6	-33.7	-58.0	-20.3	-26.6	-10.9	-4.2	4.5	33.1	-5.16
Mar	85.7	48.4	25.1	-38.2	10.5	-11.8	18.9	22.6	27.3	40.0	22.85
Apr	3.5	-24.8	-12.1	1.6	-18.7	-27.0	-12.3	-32.6	-164.9	20.8	-26.65
May	24.3	6.0	-23.3	-14.6	-5.9	-42.2	-26.5	11.2	1.9	18.6	-5.05
Jun	-5.9	26.8	21.5	153.2	-123.1	-35.4	-22.7	-31.0	23.7	72.4	7.95
Jul	124.9	22.6	-2.7	-79.0	58.7	20.4	-38.9	-28.2	-25.5	42.2	9.45
Aug	37.7	21.4	-19.9	-49.2	-0.5	34.2	15.9	-33.4	45.3	74.0	12.55
Sep	18.6	4.2	-37.1	-30.4	1.3	-76.0	-16.3	-10.6	18.1	-27.2	-15.54
Oct	10.4	1.1	-35.3	-17.6	-0.9	40.8	-31.5	-55.8	13.9	94.6	1.97
Nov	29.2	0.9	-39.4	-2.7	10.9	30.6	56.3	-7.0	43.7	39.4	16.19
Dec	27.0	11.7	-65.6	-12.9	10.8	15.4	-23.9	-25.2	-44.5	112.2	0.5

In table 7 (below) I have taken my formulae from the table of 7 point moving means: Y = 247.55 + 1.146X where X is the number of months since January 2002 and the formula from the table of least squares: Y = 244.91 + 1.192X where X is the number of months since January 1997. The calculated figures are then seasonally adjusted to give anticipated numbers of containers through the port for 2007. The figures appear to be very close.

Table 7 Year / Month			Forecast from Moving Means 247.55 + 1.146X	Average monthly Deviation	Forecast Values		Year / Month			Forecast from Least Squares 244.91 + 1.192X	Average monthly Deviation	Forecast Values
2007	Jan	120	385.07	-8.57	377		2007	Jan	120	387.95	-13.1	375
	Feb	121	386.216	-3.32	383			Feb	121	389.142	-5.2	384
	Mar	122	387.362	23.33	411			Mar	122	390.334	22.8	413
	Apr	123	388.508	-25.26	363			Apr	123	391.526	-26.7	365
	May	124	389.654	-7.34	382			May	124	392.718	-5.1	388
	Jun	125	390.8	7.17	398			Jun	125	393.91	8.0	402

A graph comparing the predicted values for 2007 from the 7 point moving means calculations and the least squares calculations is below.



Report

Description of the Method Used

I chose two moving means as when I had completed the first 5 point moving mean trend line I felt that it didn't smooth the data enough. The 7 point moving mean did smooth the series a great deal more and the trend became clearer. The disadvantage is that there are a lot more computed data points missing at the beginning and at the end of the series which makes it more difficult to obtain an overall impression of the series.

In Tables 3 and 4 (pages 5-6) the trend is calculated using the method of least squares. In the pages following this I compared the results and made predictions.

Description of the Trend

From January 2005 the moving mean line seems to trend upwards at an increasing rate to previous months. To give a quantitative figure for the trend line I entered the data into an Excel spreadsheet and produced a chart of months v data and moving averages (see Figures 1 and 2). I used the trend line option on the 7 point moving mean data. This gave me a trend line equation of:
$Y = 1.146X + 247.55$. Figure 2 on page 4 shows a graph of the original data, the 7 point moving mean and the calculated linear trend line.

There appear to be three parts to the data. Between January 1997 and January 2000 the trend is generally decreasing. From January 2000 onwards there is a general trend upwards. From January 2005 the increase is at a steeper rate. The reason for this could be the increased population growth in New Zealand since 2000, with more people needing more goods. During this time the New Zealand dollar has also been rising at a steady rate which meant we could purchase more overseas for our money. As a result there were many more imports, many of which would have come on containers.

Relevance of the forecasts

The seasonally adjusted figures for both the Moving Means data and the Least Squares data are reasonably close. The difference between the two January predictions is less than 0.5%. We can therefore be confident that the totals for the months of 2007 will lie somewhere between the two calculated figures. While the trend is upward we still have to be aware of other factors that could affect future data such as a flu epidemic, discovery of a disease such as 'foot and mouth' in New Zealand. This would have an effect on the number of export containers. A future low NZ dollar could have an effect on import containers, but higher export containers, while a high NZ dollar could have a negative effect on export containers. The data could be compared to other factors such as the value of the NZ currency or the growth of the population to see if there is any correlation.

Features of the Time Series Data

Both trend lines show an increasing number of containers moving through the port. From January 1997 until January 2005 there is a slight decreasing trend line when using the moving means. From May 2005 the trend line seems to be fairly linear and trending upwards. The rate increases from January 2005. The trend lines must be affected by the outliers. I considered outliers as any data variation of 9000 or more. e.g. July 1997 (a variation of 9443 containers), June 2000, (a variation of 16129 containers), June 2001 (a variation of -11200 containers) and April 2005 (a variation of -14514 containers).

Appropriateness of the Model

The use of the straight line least squares method may not give the most accurate result because if you drew a freehand line of best fit through the data then your line may not be truly linear. Therefore it may be more appropriate to use a parabolic or other type of trend line. When using the moving means method, it was difficult to ascertain the proper period of moving means. I initially chose 5 months because there seemed to be 5 general high/low cycles during each year - January, March, April, September and November. I then settled on the 7 month moving mean as it seemed to smooth the data better. The use of a parabolic trend line might have also given a slightly better prediction when predicting using the moving means line.

Improvements to the model

Because the trend lines are affected by the outliers, one of the methods that could be used is to replace the outliers with numbers that are closer to past values for that month. This would then have the effect of providing a closer trend line and then better predictions. One way would be to take the mean value for that month over the 10 years and replace the outlier with that value. e.g. move July 1977, from 377 to 314, June 2000 from 285 to 269 and April 2005 from 198 to 238. If using the least squares calculations you could just leave out the actual values as this method does allow for missing data. Another method would be to just take the data from January 2005 with a 4 point moving mean and calculated trend line. The April 2005 outlier would need to be adjusted.

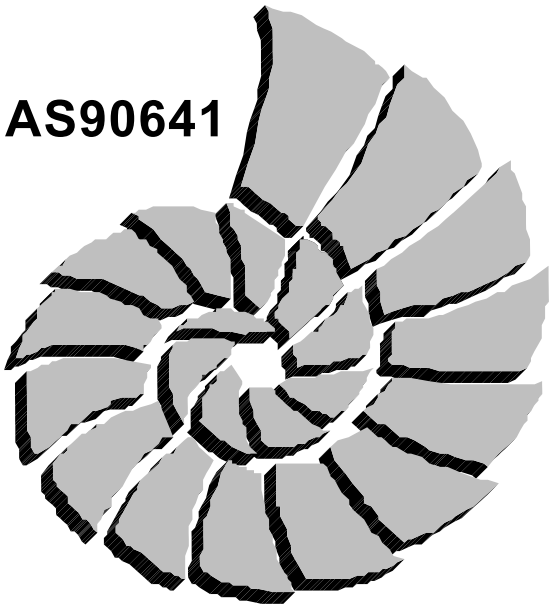
Limitations of the analysis

The predictions for 2007 assume that the conditions that gave us the data will stay the same and that the same growth will continue. Because we have a good amount of data we can be reasonably confident in our predictions over the months that follow immediately but not long term.

Exemplar provided by Katie-May Wakefield, Remuera, Auckland.

There have been slight changes to dates and the port name to the information provided in the actual assignment. This assignment received an Excellence from the examiner at Katie-May's school. All the calculations and graphs are easily replicated and Katie-May has provided some good alternatives to improve calculations and analysis.

This particular set of numbers presents students with many options. You could take the data as a whole (as Katie-May has done) with appropriate conclusions. You could also take out the outliers and give comparisons. Some assignments that used this data set still obtained excellence by providing 3 distinct trend lines and showing that each trend line had three distinct gradients. These assignments then used the data from January 2005 until December 2006 with appropriate calculations and conclusions. Note that the use of least square analysis is not necessary to obtain excellence in this Achievement Standard. Trends can be obtained with the use of Excel. While the presentation of all the data, calculations and predictions is important to achieve excellence in this Achievement Standard it is important to answer all the questions in the report fully by referring back to your calculations.



Exemplar Material (2) For

Time Series

Kim Freeman

This book provides an exemplar for
NZQA, Level 3 Mathematics
Statistics and Modelling 3.1
Determine the Trend for Time Series

Exemplar Material (2) For Time Series
Published 2007 by:
Mahobe Resources (NZ) Ltd
P.O. Box 109-760
Newmarket, Auckland 1149
New Zealand

admin@mahobe.co.nz
www.mahobe.co.nz

ISBN 0-9583564-9-1
ISBN13 9780958356497

'Exemplar Material (1) For Time Series' provides material for the New Zealand Qualifications Authority, NCEA Level 3 Mathematics Achievement Standard 3.1 of Statistics and Modelling. The standard is AS90641: Determine the trend for time series data.

Mahobe Resources are grateful for the students who gave permission to use their work as resources. The marking and grading of these was the responsibility of the teachers of these students and has not been considered by the publisher. Any alterations to the script was only to help clarify areas in which the publisher believed may assist the reader.

Mahobe Resources (NZ) Ltd welcomes feedback on the book and its usefulness.

Time Series
Internal Assessment Resource

The data set below contains quarterly figures for a retail trade series, obtained from a local Chinese takeaway, for the decade beginning March 1996 to December 2005.

Mar-96	1705		**Mar-01**	1978
Jun-96	1690		**Jun-01**	1971
Sep-96	1673		**Sep-01**	2004
Dec-96	1842		**Dec-01**	2207
Mar-97	1724		**Mar-02**	2070
Jun-97	1701		**Jun-02**	2093
Sep-97	1757		**Sep-02**	2110
Dec-97	1924		**Dec-02**	2597
Mar-98	1760		**Mar-03**	2114
Jun-98	1753		**Jun-03**	2152
Sep-98	1792		**Sep-03**	2207
Dec-98	1949		**Dec-03**	2423
Mar-99	1810		**Mar-04**	2302
Jun-99	1817		**Jun-04**	2259
Sep-99	1892		**Sep-04**	2315
Dec-99	2081		**Dec-04**	2552
Mar-00	1966		**Mar-05**	2411
Jun-00	1927		**Jun-05**	2377
Sep-00	1905		**Sep-05**	2428
Dec-00	2075		**Dec-05**	2699

1. Your task is to produce a report that includes the following:
   - time series graphs
   - analysis of the series
   - a forecast
   - appropriate conclusions

2. For the same times series write a report that presents:
   a. A table containing your series, the centred moving mean (the smooth) and individual seasonal effects (the rough).
   b. A graph of your raw data and the centred moving mean on the same axes.
   c. Describe the long term trend of your series with reasons that might account for this trend.
   d. A prediction of the value of your series for the June quarter of the year 2006. You must show how you established this result and state any assumptions you have made.
   e. The equation(s) for your model together with a justification for your choice of model. You may choose to describe your data by other than a single linear trend over the entire range.
   f. A discussion of at least two of:
      i. The relevance and usefulness of the June 2006 forecast
      ii. Identify and comment on any further features of the time-series data in addition to the trend.
      iii. The seasonal effects.
      iv. Possible limitations to your model.
      v. Limitations of your analysis and any potential sources of bias.

Table 1. Retail Sales for Chinese Takeaway data set along with 4 month moving average (MA), centred moving average, individual monthly variations and average monthly variations (seasonal effect).

		$	MA	CMA	Variation
Mar-96	1	1705			
Jun-96	2	1690			
Sep-96	3	1673	1727.5	1729.875	-56.875
Dec-96	4	1842	1732.25	1733.625	108.375
Mar-97	5	1724	1735	1745.5	-21.5
Jun-97	6	1701	1756	1766.25	-65.25
Sep-97	7	1757	1776.5	1781	-24
Dec-97	8	1924	1785.5	1792	132
Mar-98	9	1760	1798.5	1802.875	-42.875
Jun-98	10	1753	1807.25	1810.375	-57.375
Sep-98	11	1792	1813.5	1819.75	-27.75
Dec-98	12	1949	1826	1834	115
Mar-99	13	1810	1842	1854.5	-44.5
Jun-99	14	1817	1867	1883.5	-66.5
Sep-99	15	1892	1900	1919.5	-27.5
Dec-99	16	2081	1939	1952.75	128.25
Mar-00	17	1966	1966.5	1968.125	-2.125
Jun-00	18	1927	1969.75	1969	-42
Sep-00	19	1905	1968.25	1969.75	-64.75
Dec-00	20	2075	1971.25	1976.75	98.25
Mar-01	21	1978	1982.25	1994.625	-16.625
Jun-01	22	1971	2007	2023.5	-52.5
Sep-01	23	2004	2040	2051.5	-47.5
Dec-01	24	2207	2063	2078.25	128.75
Mar-02	25	2070	2093.5	2106.75	-36.75
Jun-02	26	2093	2120	2168.75	-75.75
Sep-02	27	2110	2217.5	2223	-113
Dec-02	28	2597	2228.5	2235.875	361.125
Mar-03	29	2114	2243.25	2255.375	-141.375
Jun-03	30	2152	2267.5	2245.75	-93.75
Sep-03	31	2207	2224	2247.5	-40.5
Dec-03	32	2423	2271	2284.375	138.625
Mar-04	33	2302	2297.75	2311.25	-9.25
Jun-04	34	2259	2324.75	2340.875	-81.875
Sep-04	35	2315	2357	2370.625	-55.625
Dec-04	36	2552	2384.25	2399	153
Mar-05	37	2411	2413.75	2427.875	-16.875
Jun-05	38	2377	2442	2460.375	-83.375
Sep-05	39	2428	2478.75		
Dec-05	40	2699			
Mar-06	41				
Jun-06	42				

	Mar	Jun	Sep	Dec
1996			-56.9	108.4
1997	-21.5	-65.3	-24	132
1998	-42.9	-57.4	-27.8	115
1999	-44.5	-66.5	-27.5	128.3
2000	-2.13	-42	-64.8	98.25
2001	-16.6	-52.5	-47.5	128.8
2002	-36.8	-75.8	-113	361.1
2003	-141.0	-93.8	-40.5	138.6
2004	-9.25	-81.9	-55.6	153
2005	-16.9	-83.4		
Seasonal Effect	**-36.9**	**-68.7**	**-50.8**	**151.5**

Figure 1. Retail Sales for the Chinese Takeaway data set. Actual values with centred moving average (CMA) and calculated linear trend line (Excel): y = 21.299x + 1584.7

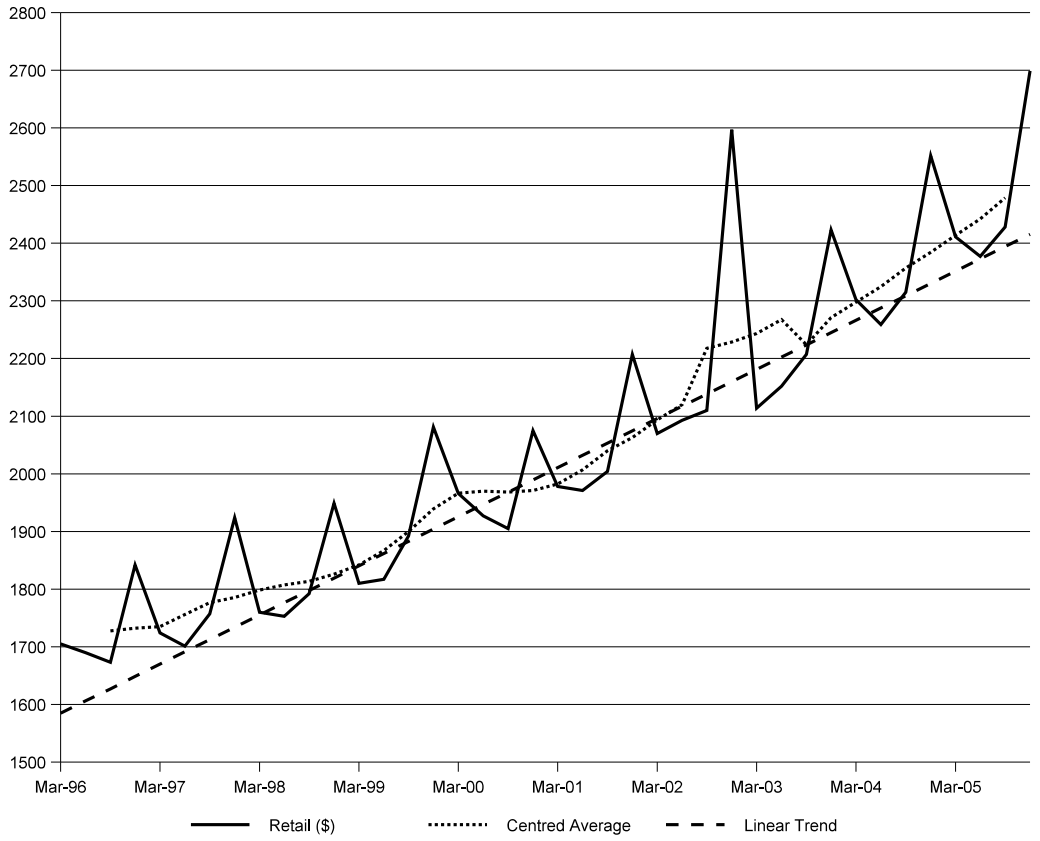


Figure 2. Retail Sales for Chinese Takeaway. Actual values with centred moving average (CMA) and parabolic trend curve (Excel): $y = 0.2478x^2 + 10.646x + 1672.5$.

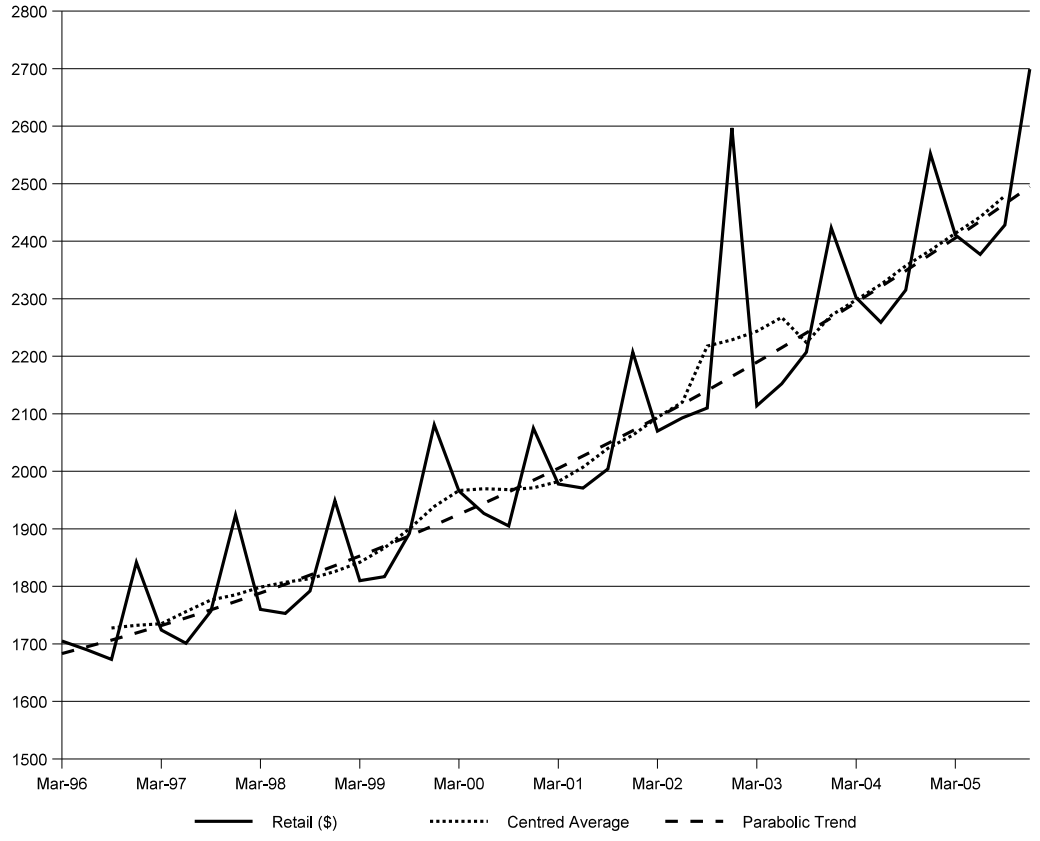


Table 2. Retail Sales for Chinese Takeaway data set with one outlier changed. Table also includes the new 4 month moving averages (MA), centred averages, individual monthly variations and average monthly variations (seasonal effect).

		$	MA	CMA	Variation
Mar-96	1	1705			
Jun-96	2	1690			
Sep-96	3	1673	1727.5	1729.875	-56.875
Dec-96	4	1842	1732.25	1733.625	108.375
Mar-97	5	1724	1735	1745.5	-21.5
Jun-97	6	1701	1756	1766.25	-65.25
Sep-97	7	1757	1776.5	1781	-24
Dec-97	8	1924	1785.5	1792	132
Mar-98	9	1760	1798.5	1802.875	-42.875
Jun-98	10	1753	1807.25	1810.375	-57.375
Sep-98	11	1792	1813.5	1819.75	-27.75
Dec-98	12	1949	1826	1834	115
Mar-99	13	1810	1842	1854.5	-44.5
Jun-99	14	1817	1867	1883.5	-66.5
Sep-99	15	1892	1900	1919.5	-27.5
Dec-99	16	2081	1939	1952.75	128.25
Mar-00	17	1966	1966.5	1968.125	-2.125
Jun-00	18	1927	1969.75	1969	-42
Sep-00	19	1905	1968.25	1969.75	-64.75
Dec-00	20	2075	1971.25	1976.75	98.25
Mar-01	21	1978	1982.25	1994.625	-16.625
Jun-01	22	1971	2007	2023.5	-52.5
Sep-01	23	2004	2040	2051.5	-47.5
Dec-01	24	2207	2063	2078.25	128.75
Mar-02	25	2070	2093.5	2106.75	-36.75
Jun-02	26	2093	2120	2137.875	-44.875
Sep-02	27	2110	2155.75	2161.25	-51.25
Dec-02	28	2350	2166.75	2174.125	175.875
Mar-03	29	2114	2181.5	2193.625	-79.625
Jun-03	30	2152	2205.75	2214.875	-62.875
Sep-03	31	2207	2224	2247.5	-40.5
Dec-03	32	2423	2271	2284.375	138.625
Mar-04	33	2302	2297.75	2311.25	-9.25
Jun-04	34	2259	2324.75	2340.875	-81.875
Sep-04	35	2315	2357	2370.625	-55.625
Dec-04	36	2552	2384.25	2399	153
Mar-05	37	2411	2413.75	2427.875	-16.875
Jun-05	38	2377	2442	2460.375	-83.375
Sep-05	39	2428	2478.75		
Dec-05	40	2699			
Mar-06	41				
Jun-06	42				

	Mar	Jun	Sep	Dec
1996			-56.9	108.4
1997	-21.5	-65.3	-24	132
1998	-42.9	-57.4	-27.8	115
1999	-44.5	-66.5	-27.5	128.3
2000	-2.13	-42	-64.8	98.25
2001	-16.6	-52.5	-47.5	128.8
2002	-36.8	-44.9	-51.3	175.9
2003	-79.6	-62.9	-40.5	138.6
2004	-9.25	-81.9	-55.6	153
2005	-16.9	-83.4		
Seasonal Effect	**-30**	**-61.8**	**-44**	**130.9**

Figure 3. Retail Sales for Chinese Takeaway data set with Dec 02 outlier changed (recalculated from 2597 to 2350). Actual values with centred average (CMA) and calculated linear trend line (Excel): y = 20.823x + 1588.1

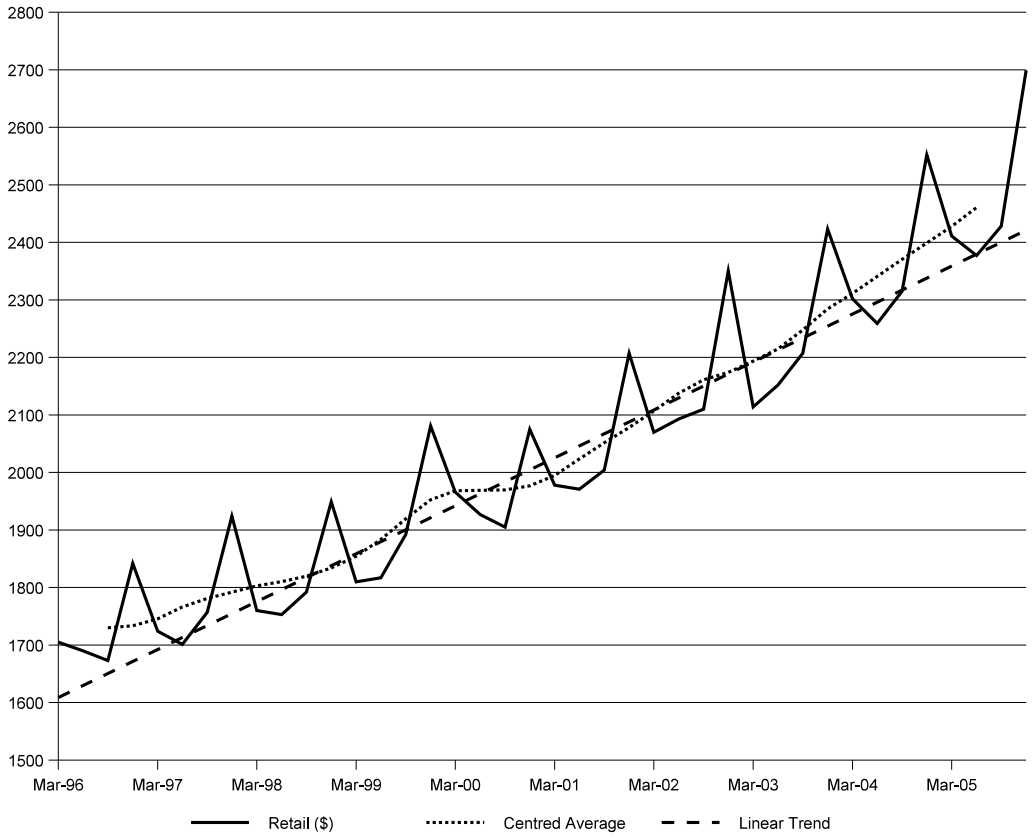


Table 3. Retail Sales for Chinese Takeaway data set from March 96 to Jun 01 along with 4 month moving average (MA), centred moving average, individual monthly variations and average monthly variations (seasonal effect).

		$	MA	CA	Variation
Mar-96	1	1705			
Jun-96	2	1690			
Sep-96	3	1673	1727.5	1729.875	-56.875
Dec-96	4	1842	1732.25	1733.625	108.375
Mar-97	5	1724	1735	1745.5	-21.5
Jun-97	6	1701	1756	1766.25	-65.25
Sep-97	7	1757	1776.5	1781	-24
Dec-97	8	1924	1785.5	1792	132
Mar-98	9	1760	1798.5	1802.875	-42.875
Jun-98	10	1753	1807.25	1810.375	-57.375
Sep-98	11	1792	1813.5	1819.75	-27.75
Dec-98	12	1949	1826	1834	115
Mar-99	13	1810	1842	1854.5	-44.5
Jun-99	14	1817	1867	1883.5	-66.5
Sep-99	15	1892	1900	1919.5	-27.5
Dec-99	16	2081	1939	1952.75	128.25
Mar-00	17	1966	1966.5	1968.125	-2.125
Jun-00	18	1927	1969.75	1969	-42
Sep-00	19	1905	1968.25	1969.75	-64.75
Dec-00	20	2075	1971.25	1976.75	98.25
Mar-01	21	1978	1982.25		
Jun-01	22	1971			

	Mar	Jun	Sep	Dec
1996			-56.9	108.4
1997	-21.5	-65.3	-24	132
1998	-42.9	-57.4	-27.8	115
1999	-44.5	-66.5	-27.5	128.3
2000	-2.13	-42	-64.8	98.25
Seasonal Effect	**-27.8**	**-57.8**	**-40.2**	**116.4**

Figure 4. Retail Sales for Chinese Takeaway - March 96 to Jun 01. Actual values with centred moving average (CMA) and calculated linear trend line y = 16.36x + 16.46

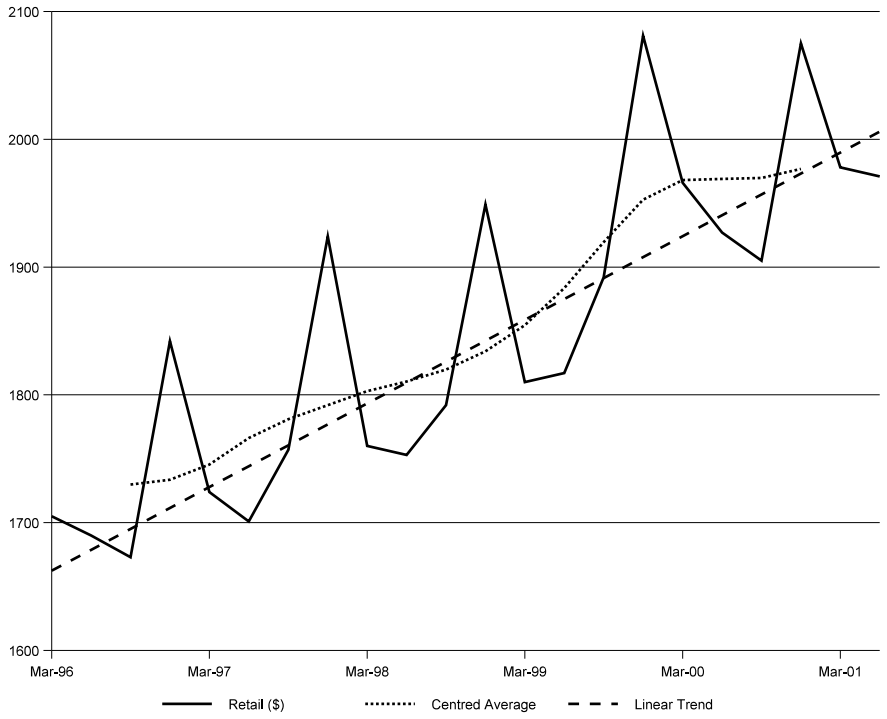
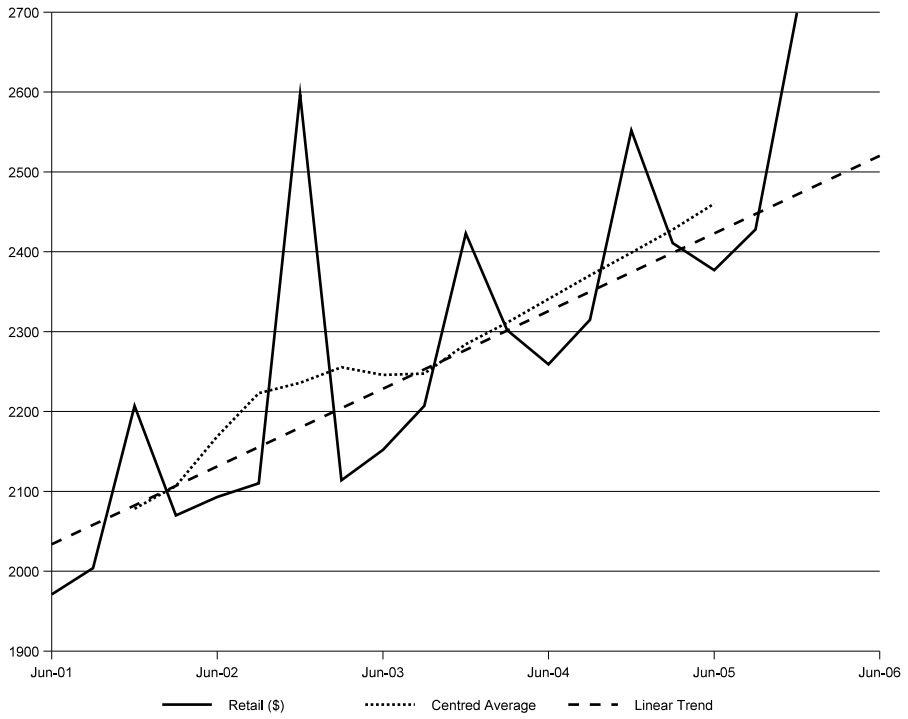


Table 4. Retail Sales for Chinese Takeaway from Jun 01 to Jun 06 along with 4 month moving average (MA), centred average, individual monthly variations and average monthly variations (seasonal effect).

		$	MA	CA	Variation
Jun-01	1	1971			
Sep-01	2	2004			
Dec-01	3	2207	2063	2078.25	128.75
Mar-02	4	2070	2093.5	2106.75	-36.75
Jun-02	5	2093	2120	2168.75	-75.75
Sep-02	6	2110	2217.5	2223	-113
Dec02	7	2597	2228.5	2235.875	361.125
Mar-03	8	2114	2243.25	2255.375	-141.375
Jun-03	9	2152	2267.5	2245.75	-93.75
Sep-03	10	2207	2224	2247.5	-40.5
Dec-03	11	2423	2271	2284.375	138.625
Mar-04	12	2302	2297.75	2311.25	-9.25
Jun-04	13	2259	2324.75	2340.875	-81.875
Sep-04	14	2315	2357	2370.625	-55.625
Dec-04	15	2552	2384.25	2399	153
Mar-05	16	2411	2413.75	2427.875	-16.875
Jun-05	17	2377	2442	2460.375	-83.375
Sep-05	18	2428	2478.75		
Dec-05	19	2699			
Mar-06	20				
Jun-06	21				

	Mar	Jun	Sep	Dec
2001				128.8
2002	-36.8	-75.8	-113	361.1
2003	-141	-93.8	-40.5	138.6
2004	-9.25	-81.9	-55.6	153
2005	-16.9	-83.4		
Seasonal Effect	**-51.1**	**-83.7**	**-69.7**	**195.4**

Figure 5. Retail Sales for Chinese Takeaway - Jun 01 to Jun 06
Actual values with centred average (CA) and calculated linear trend line y = 24.317x + 2009.6



Report

The Chinese takeaway sales are increasing at a rate of $21.30 per quarter (based on the trend line gradient - see Figure 1, page 3). Over a 3 month period the figure is insignificant however does show that there is growth. Growth in takeaway sales may be due to increased customers, good food, good marketing or word of mouth from satisfied customers.

I predict that the sales in June 2006 will be $2410.55.
This number was obtained using the linear trend line equation of y = 21.299x + 1584.7 and includes the seasonal effect for June. This value assumes the trend will continue and that sales will have a linear growth of $21.30 per quarter year.

An improvement to the prediction is to make the trend line a parabolic curve as the sales trend fits a model which is similar to a parabola. However, as with many models, this may be valid for a limited number of years. Using the parabolic trend line with the equation $y = 0.2478x^2 + 10.646x + 1672.5$ the prediction for June 2006 would be $2488.04. The parabolic trend is illustrated on Figure 2, page 3.

There seems to be an outlier in December 2002 which may have dragged the trend line upwards and resulted in higher than usual prediction. To improve this I removed the outlier to a more realistic amount of $2350. From this a new trend equation was formed: y = 20.823x + 1588.1 and a new prediction of $2400.82. The calculations are shown on Table 2 page 4 and illustrated on Figure 3, page 5. This means that the sales would increase at $20.82 per quarter.

Although the predictions seem reasonable when compared to previous quarters the trend appears to vary through 1996 to 2005, with a more apparent change from June 2001. The original calculations would have smoothed any upward trend between these years and could result in misleading calculations and predictions. Therefore, to improve the predictions I separated the data into two separate time periods - one before June 2001 and one after June 2001. This resulted in a new trend line equation of y = 24.317x + 2009.6 for March 2002 onwards - an increase of almost $8 per quarter on the March 96 to June 2001 calculated trend line (which was y = 16.36x + 16.46). This gives a prediction of $2436.56 for June 2006. The difference between the original and this prediction is only $26. The calculations for these are found on Tables 3 and 4 and are illustrated in Figures 4 and 5 on pages 6 and 7. Note - the parabolic model is over $51 higher.

Discussion: 1. Further features of the time series data (seasonal effects)
The seasonal effect for the December quarter is positive. This shows that sales in December were above the centred average by $151.49. The seasonal effects for March, June and September were all negative. This indicated that they were below the centred average by $36.88, $68.71 and $50.83 (see table 1). The variations in the March and September quarters fluctuate each year with inconsistent ranges. In the June and December quarters, variations increase at a steady rate from 2000 onwards. There is a significant change from 2001 to 2002 for the December quarters with an almost three fold increase ($128.75 to $361.13) and then a substantial decrease for the following year ($361.13 to $138.63). Similarly from 1999 to 2000 (for the June quarter) there is a decrease in variation by $24.50.

Discussion: 2. Possible limitations to the models
All of the trend calculations give similar results for June 2006. These range between $2410 and $2486. Longer term predictions may be misleading as there may be external factors that change the actual results such as a decrease in customers, food quality deteriorates or change of ownership.

Exemplar provided by William Lee, St Heliers, Auckland.
This assignment received an Excellence from the examiner at William's school.