June 2014 The deputy principal of a college kept a daily record of the percentage of students absent. The data for weeks 3 and 4 of the college year are shown in the table below, together with the values of an appropriate moving average.

Week	Day	Percentage absent	Moving average
3	Monday	6.4	
	Tuesday	4.6	
	Wednesday	4.1	6.40
	Thursday	6.5	m
	Friday	10.4	7.88
4	Monday	10.1	8.58
	Tuesday	8.3	9.32
	Wednesday	7.6	10.00
	Thursday	10.2	
	Friday	13.8	

(a) Calculate the value of the missing moving average, m.

[3 marks]

(b) The values of the percentages absent are plotted on **Figure 1** opposite.

(i) Plot the moving averages on **Figure 1** and draw a trend line by eye.

[2 marks]

(ii) Hence describe the variation and trend in percentage absent.

[2 marks]

(c) Assuming that the current pattern continued, use the trend line and the seasonal effect for Friday to forecast the percentage absent on the Friday of week 5, showing how you obtained this forecast.

[5 marks]

(d) The deputy principal was keen that the current pattern should not continue and so, at the beginning of week 5, introduced incentives to encourage attendance. The percentage absent on the Friday of week 5 was 15.2 .

Make **two** comments about the success of the deputy principal's incentives.

[2 marks]

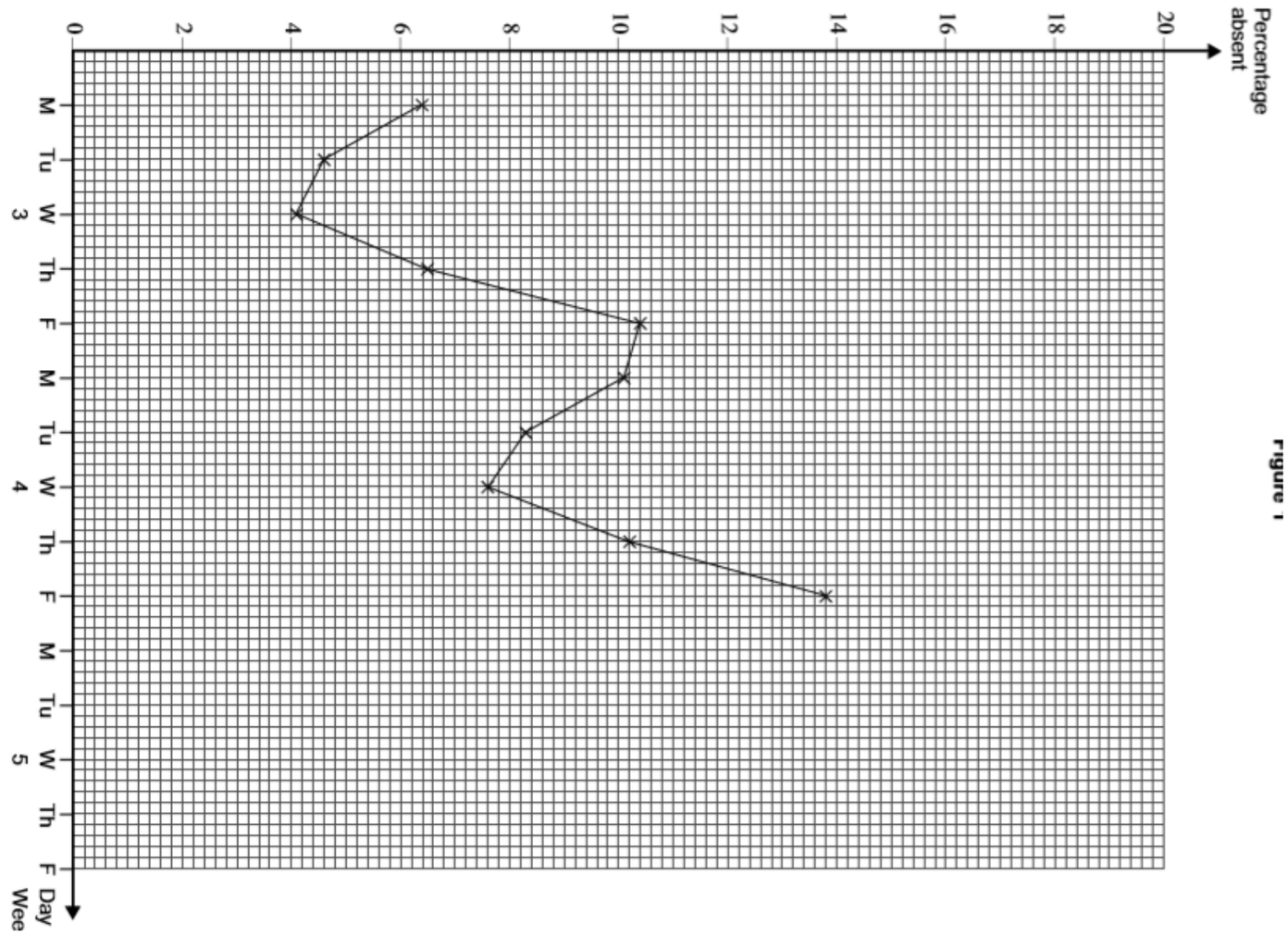


Figure 1

June 2016

The heights of a sample of 240 female students and 240 male students were measured.

The data for the female students are summarised as a box plot in **Figure 1**.

The data for the male students are summarised as a cumulative frequency graph **Figure 2**.

Using the information in the two figures, compare the distribution of heights for th female students with that for the male students. You should make reference to the difference, if any, between:

(a) the average values of height;

(b) the values of a measure of spread;

(c) the symmetry, or otherwise, of the two distributions.

[6 marks]

Figure 1

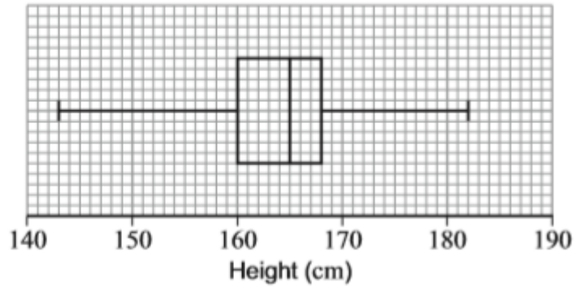
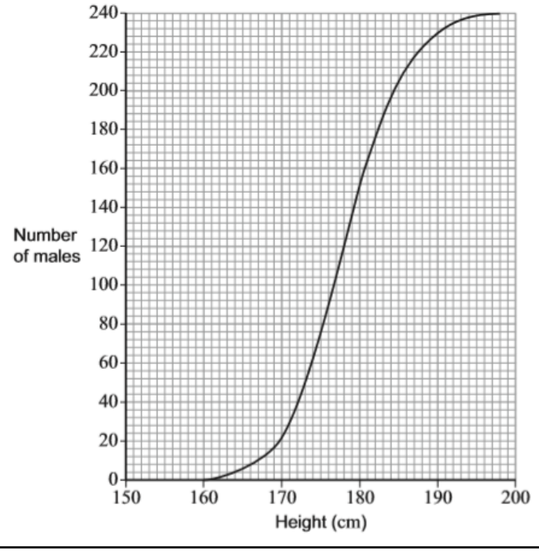


Figure 2



UNIVARIATE DIAGRAMS
Revision Mat

Time Series

A Time Series Graph is the result of recording a variable at intervals of _____

Time series are analysed so that they may be _____

_____

BE AWARE: Forecasting beyond the data may be unreliable because _____

_____

Trend

Trend is used to describe

_____

The 'trend' line should be a long term smooth movement showing an _____

or _____

WE DO NOT SAY A TREND HAS A POSITIVE OR NEGATIVE CORRELATION – THIS IS FOR BIVARIATE DATA ON A SCATTER GRAPH

Variation

Seasonal Variation:_____

_____

Random Variation:_____

_____

Short-Term Variation:_____

_____

Moving Averages

Moving averages are an efficient and practical way of finding the trend and mean you make more

_____ predictions

Moving averages tend to reduce the amount of _____ present in a time series

An 'n-point' moving average relies on how many _____ are repeated

For example:

- If a year is split into 4 quarters we have a _____ point moving average
- If a week is split into 5 days we have a _____ point moving average

To calculate a moving average we work out the _____ value for each 'n' season

This will include overlapping values as we 'move' along in time

Seasonal Effect

Seasonal Effect is the difference between the actual value and the value read from the trend line at a given point (season)

Seasonal Effect = ☐

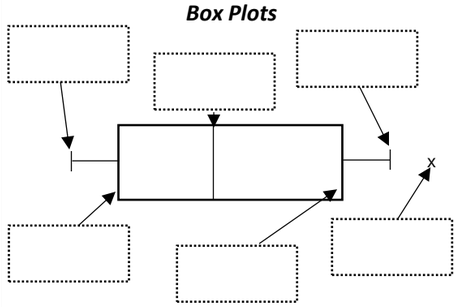
The average of these seasonal effects can be calculated for one season over several years to find the

_____

To then forecast into the future, we add this to an estimated value read from the

_____

Box Plots



Comparing Box Plots

SP _____

_____

A _____

_____

S _____

_____

I _____

_____

Misleading Statistics

Statistics are often used and manipulated to make the data look better (or worse) without lying to the reader.

When looking at any statistics we must look for

What is 'misleading'	What impact this may have on the reader
Missing labels	
Unequal gaps along scales	
Large breaks in the scales	
3D graphs	
Incorrectly calculated angles	
Time series without repeated seasons	
Different scales used for box plots	
Key not included	

The Use of Software

For this new topic you are expected to be able to:

- Read and interpret data presented in a spreadsheet or database.
- Understand and use the terms: filter, sort, query and field.

Filter: _____

_____

Sort: _____

_____

Query: _____

_____

Field: _____

_____

- Describe methods for extracting specific information from a spreadsheet or database using the correct terms described above.

Remember key things that a computer can do much more conveniently that a person:

- _____
- _____
- _____

Statistical Diagrams

Univariate diagrams are those which represent one variable at a time

They can be used to represent all types of data:

Qualitative: _____

Discrete: _____

Continuous: _____

For each of the univariate diagrams below, select what sort of data they can be used for

N.B. most diagrams can be used for more than one type of data

	Qualitative	Discrete	Continuous
Pictogram			
Bar Chart			
Vertical Line Graph			
Pie Chart			
Tally Chart			
Cumulative Frequency Diagram			
Box and Whisker Plot			
Histogram			
Stem and Leaf Diagram			
Time Series			
Scatter Graph			