Full Name: SOLUTIONS

MATHEMATICS APPLICATIONS

Test 1 – Data Analysis and Linear Models

Chapters 1 and 2

Semester 1 2018

Section One - Calculator Free

Time allowed for this section
Working time for this section: 25 minutes
Marks available: 27 marks

Material required/recommended for this section
To be provided by the supervisor
This Question/Answer booklet
Formula sheet

To be provided by the candidate
Standard items: pens, pencils, pencil sharpener, eraser, correction fluid, ruler, highlighters

Special items: Nil

Important note to candidates
No other items may be used in this section of the examination. It is **your** responsibility to ensure that you do not have any unauthorised notes or other items of a non-personal nature in the examination room. If you have any unauthorised material with you, hand it to the supervisor **before** reading any further.

1. (7 marks)
   Before a fitness campaign at a high school, 50 students were chosen at random from each year group and asked the following questions:

   Question 1:   Which one of the following modes of transport do you use to travel to and from school?

   Category A: walking/cycling
   Category B: public transport
   Category C: private car

   Question 2:   Which year group are you in?

   The campaign organisers wished to determine whether age group affected the students' likelihood of walking/cycling to and from school.

   The results of the survey are shown in the table below.

	Category A	Category B	Category C	Total
Year 7	19	11 ✓	20	50
Year 8	12	17	21	50
Year 9	13	14	23	50
Year 10	11	18 ✓	21	50 ✓ $\left(\frac{1}{2} \text{ each}\right)$
Year 11	10	15	25	50
Year 12	8	17	25	50
Total	73	92 ✓	135	300

   a. Complete the missing entries in the table above.   [2]

   b. Compare the percentages of students in Year 7 and Year 12 who use Category A as a
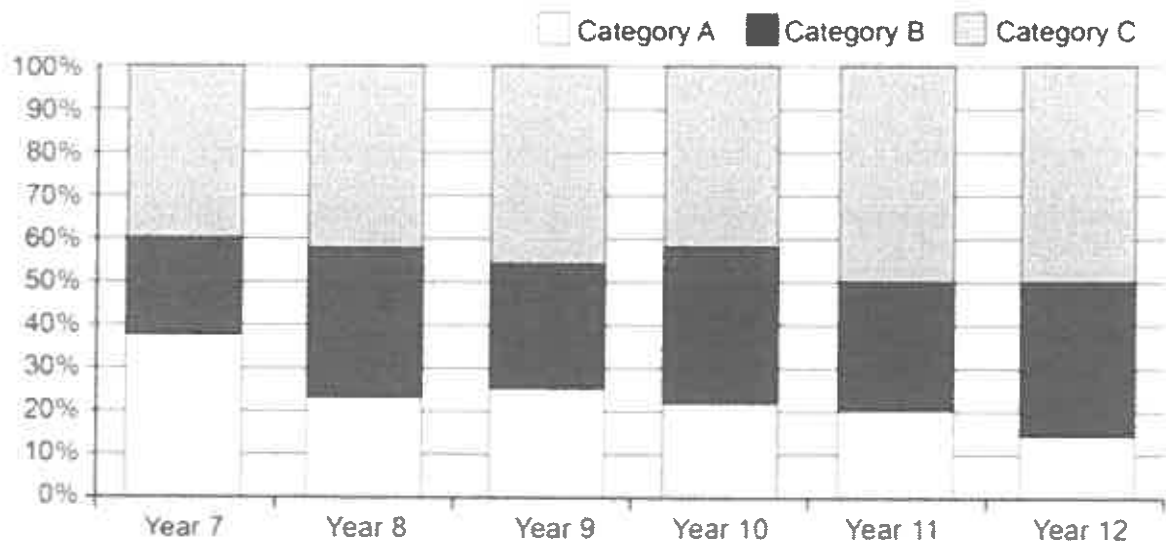   mode of transport and comment on your results.   [2]

   38% of Yr 7's use CAT A   ✓
   16% of Yr 12's use CAT A

   Significantly less Yr 12's use CATA. Could be
   because of driving/independence. Yr 7's still
   more active.
                                    ✓ viable reasoning

The data given in the table in part a. have been displayed as a divided column graph below.

c. Using the graph above or another method, comment on:

    i.   the association between 'Year group' and 'Category A'.   [1]

There appears to be a decrease in the use of Cat A as students move through the years.

    ii.   the association between 'Year group' and 'Category C'.   [1]

Use of Cat C increases as students move through the years.

    iii.   the association between 'Category A' and 'Category B and C combined'.   [1]

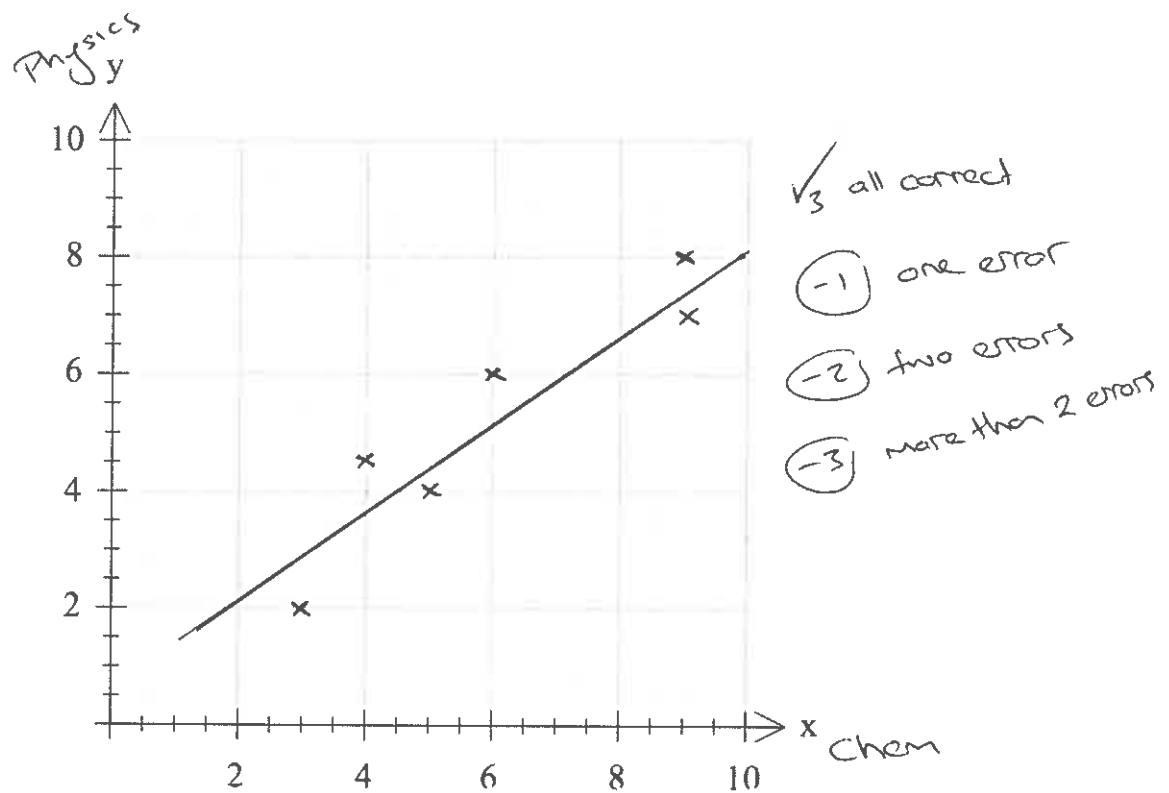*The use of motorised transport increases over time*

2. (15 marks)
   The marks for Chemistry and Physics tests (out of 10) were recorded in the table below for 7 students:

Chemistry (x)	3	4	5	6	7	9	9
Physics (y)	2	4.5	4	6	Absent	7	8

a) Draw a scatterplot to represent this information.   [3]

b) Describe the strength and direction of the association.   [2]

Strong positive association ✓✓

c) Draw the line of best fit 'by eye' and hence estimate a Physics mark for the student who was unable to sit that test (to the nearest half mark).   [2]

accept a variety of valid lines ✓ and an estimate based on line. ✓

My line estimate is 5.8

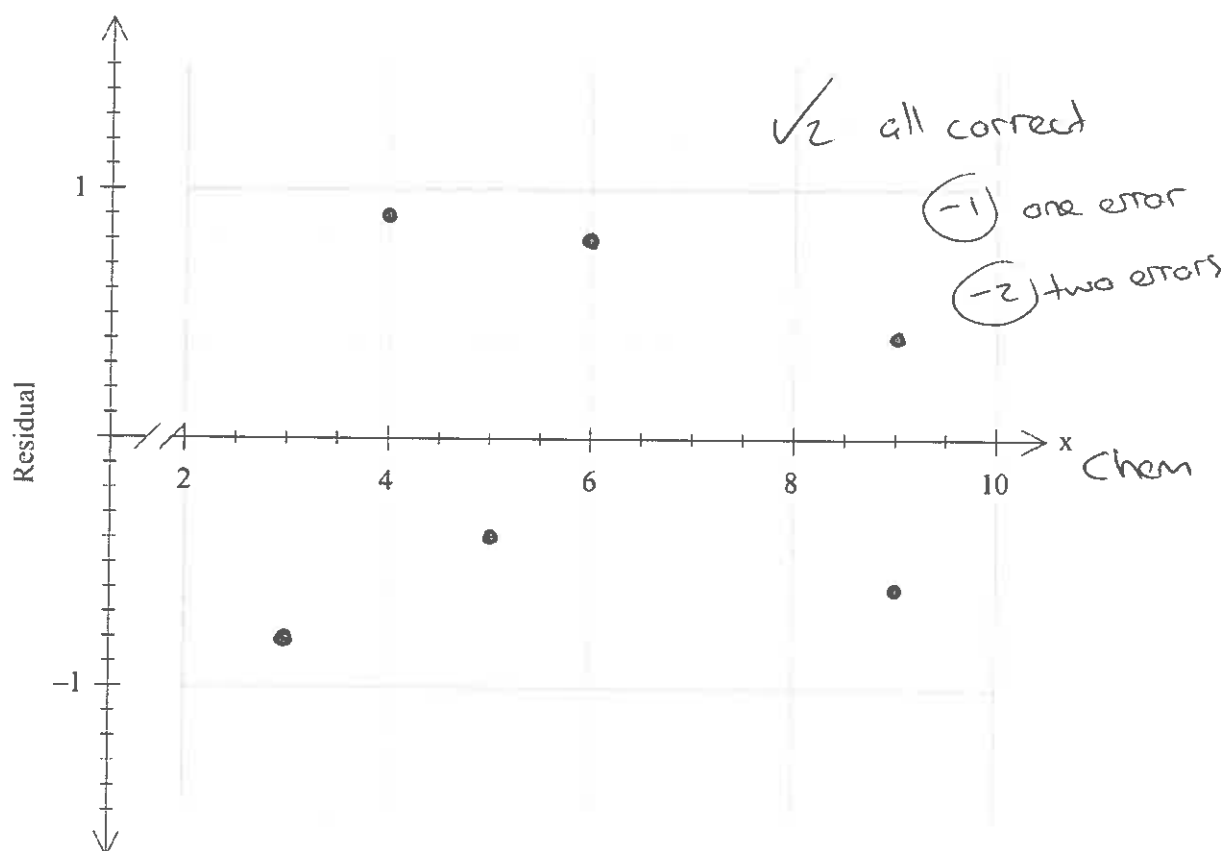The 'Line of Best Fit' equation for the association is: y = 0.8x + 0.4

d) Complete the table of residuals to 1 d.p. [4]

Chemistry (x)	3	4	5	6	9	9
Physics (y)	2	4.5	4	6	7	8
Predicted Physics	2.8	3.6	4.4	5.2	7.6	7.6
Residual	-0.8	0.9	-0.4	0.8	-0.6	0.4

✓ ✓ ✓ ✓

e) Draw a residual plot. [2]

√₂ all correct

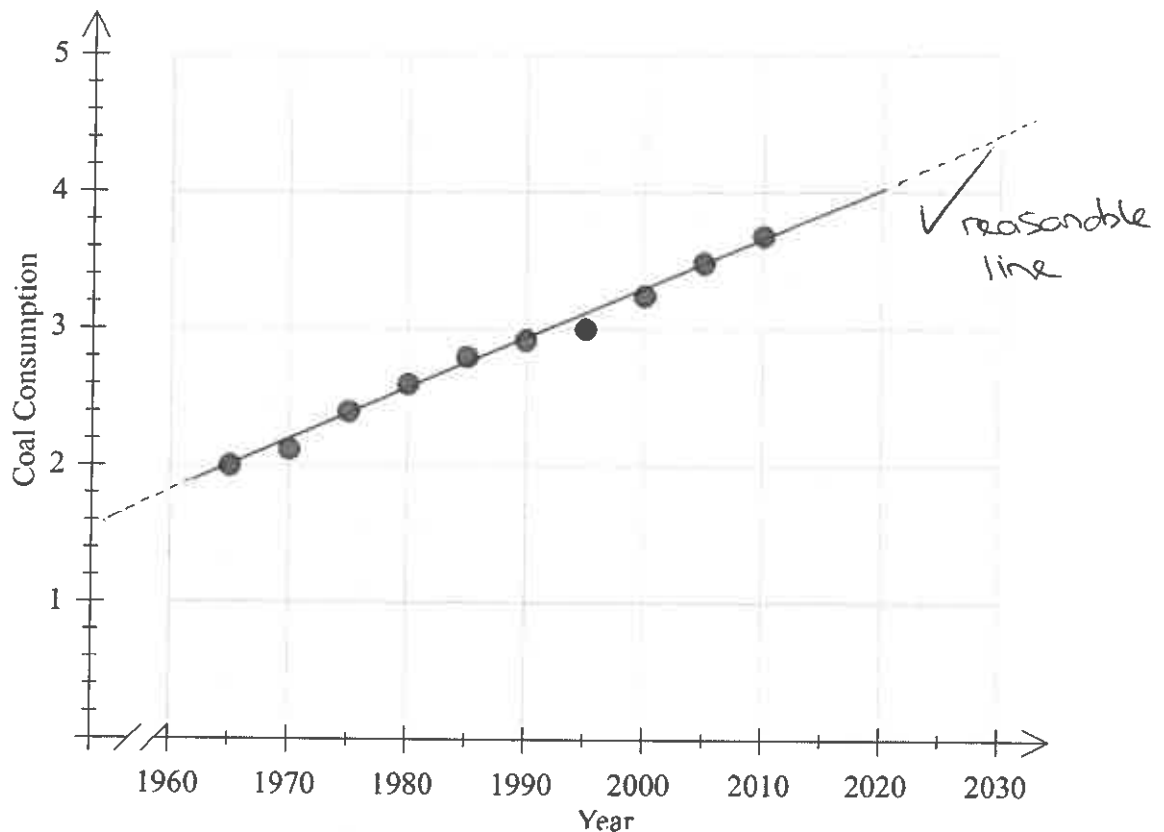(-1) one error

(-2) two errors

x Chem

Residual

f) Using the residual plot as a reference, decide if the data being investigated is linear or non-linear. [2]

Pattern appears random ✓ therefore the data is likely linear ✓

3. (5 marks)
    The worldwide consumption of coal (in billion tonnes of oil equivalent, BTOE) is shown in the graph below from 1965 until 2010.

a) Add a trend line, by eye, to the scatterplot.   [1]

b) Estimate the worldwide consumption of coal in

    (i) 1960.   1.8 billion tonnes   ✓   [1]
                                                              based on
    (ii) 2030.   4.4 billion tonnes   ✓   line.   [1]

c) Which estimate in (c) is more reliable? Explain your reasoning.   [2]

        (i) 1960 is more reliable. ✓
            still extrapolation but closer to the
            data set ✓

End of Section One

Full Name: __SOLUTIONS__

MATHEMATICS
APPLICATIONS

Test 1 – Data Analysis and Linear Models

Chapters 1 and 2

Semester 1 2018

Section Two - Calculator Assumed

Time allowed for this section
Working time for this section: 30 minutes
Marks available: 34 marks

Material required/recommended for this section
To be provided by the supervisor
This Question/Answer booklet
Formula sheet

To be provided by the candidate
Standard items: pens, pencils, pencil sharpener, eraser, correction fluid, ruler, highlighters

Special items: drawing instruments, templates, notes on one unfolded sheet of A4 paper, and up to three calculators satisfying the conditions set by the Curriculum Council for this course.

Important note to candidates
No other items may be used in this section of the examination. It is **your** responsibility to ensure that you do not have any unauthorised notes or other items of a non-personal nature in the examination room. If you have any unauthorised material with you, hand it to the supervisor **before** reading any further.

4. (8 marks)

Data was gathered on the number of items sold during the first 20 days of November at an etsy store.

| 14 | 22 | 13 | 17 | 29 | 4 | 17 | 14 | 21 | 19 |
| 17 | 23 | 45 | 5 | 9 | 14 | 17 | 25 | 18 | 11 |

a. Calculate the **two** measures of centre and **two** measures of spread. [2]

Measures of centre

Mean   17.1

Mode   17

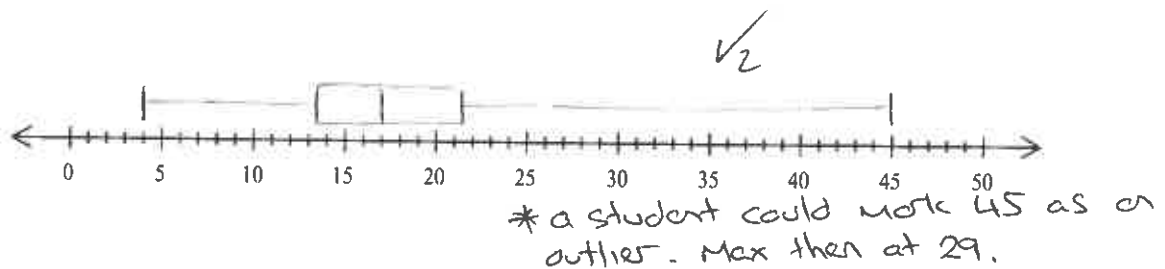Median   17

2 of these
(1)

Measures of spread

Range   41

IQR   8

S.D.   8.7 $\left(\underset{s}{8.9}\right)$ either
    σ

Var   75.7 $\left(\underset{s^2}{79.2}\right)$ either
    σ^2

2 of these
(1)

b. Draw a boxplot that represents this data. [2]

✓/2

* a student could mark 45 as an outlier. Max then at 29.

c. Are there any possible outliers? Explain [2]

no calculation expected

45 appears to be an outlier. well removed ✓
from the dataset. ✓

d. Describe the shape of the data [1]

Skewed to the left. (Negatively skewed)
✓

e. What percentage of daily sales are over 20 items? [1]

6 over 20

$\frac{6}{20} \times 100 = 30\%$ ✓

5. (8 marks)

The accompanying table shows the different makes of cars parked at three different suburban shopping centres on a school-day morning. The shopping centres A, B and C are located respectively at high, middle and low income suburbs.

	A	B	C
Australia	70	60	40
German	150	80	20
Korean	40	130	120
Japanese	90	140	110
Others	60	70	50

a. Complete the table below showing the row percentages. [2]

	A	B	C
Australia	41	35	24
German	60 ✓	32 ✓	8 ✓
Korean	14	45	41
Japanese	26	41	33 ✓
Others	33	39	28

$\left(\frac{1}{2} each\right)$

b. Complete the table below showing the column percentages. [2]

	A	B	C
Australia	17	13	12
German	37 ✓	17	6
Korean	10 ✓	27	35 ✓
Japanese	22	29	32 ✓
Others	15	15	15

$\left(\frac{1}{2} each\right)$

c. Determine with reasons if there is an association between the make of cars parked and the level of income of the suburb. Clearly identify the response and explanatory variables. [4]

explanatory: car make ✓
response: level of income ✓

German cars appear more popular in high income ✓ suburbs

Korean cars more popular in middle-low income suburbs ✓

So an association appears to exist.