education equals

Topic: Associations between two numerical variables
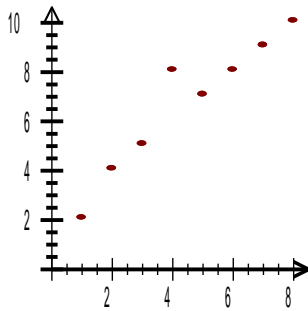
Time: 45 mins   Marks: /45 marks
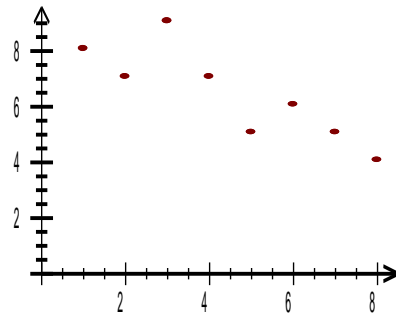
Calculator Assumed

Question One: [4 marks]

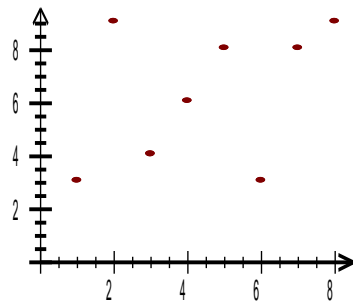Match the correlation coefficient to the correct scatter plot.
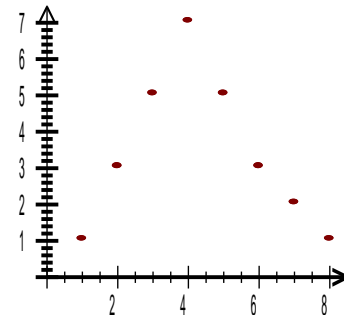
A

B

C

D

Correlation coefficient (r)	−0.84	0.40	0.95	−0.18
Scatter plot				

Question Two: [2, 2, 2, 2: 8 marks]

Describe the likely relationship between each of pair of variables.

a)   Height vs the weight of maths Applications students in your class.

b)   Value of a 2015 Holden Commadore vs time.

c)   Fuel consumption of a vehicle vs the kilometers it has travelled.

d)   Hours spent studying for maths Applications vs the test results obtained.

Question Three: [1, 1, 1: 3 marks]

Match the likely correlation coefficient, in the table at the bottom of the page, with the two variables described.

a)   Alcohol consumption vs liver function.

b)   Birth rate (number of births per 1000 population) vs female life expectancy.

c)   Energy (Kj) vs fat content (g).

$r = -0.89$	$r = 0.56$	$r = 0.90$	$r = -0.70$

Question Four: [2, 3: 5 marks]

Some people feel that it is more difficult to sleep when temperatures are higher. It has been claimed that, "High overnight temperatures cause people to sleep less than the standard 10 hours."

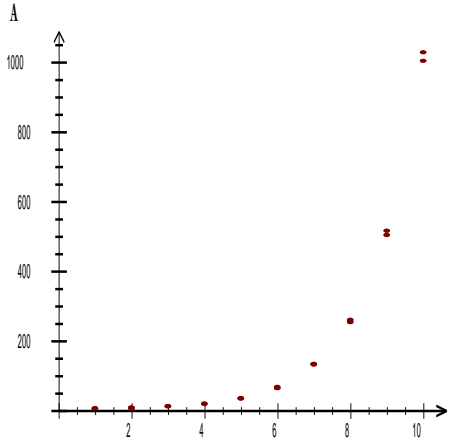a)   State the explanatory and the response variable in the above claim.

The table below shows the hours slept and the maximum temperature reached overnight for a sample of 18 people.

Max Temperature Over Night	Number of Hours slept
10	9
14	8
11	9
12	9
9	13
20	6
15	7
15	6
16	8
17	6
8	12
11	7
21	6
19	5
18	6
18	7
13	7

b)   Given the above data, comment on the claim made. Consider the use of statistical measures to support your answer where necessary.

Question Five: [3, 2: 5 marks]

a)   Order the following scatter graphs from strongest positive linear relationship to
        weakest negative linear relationship.
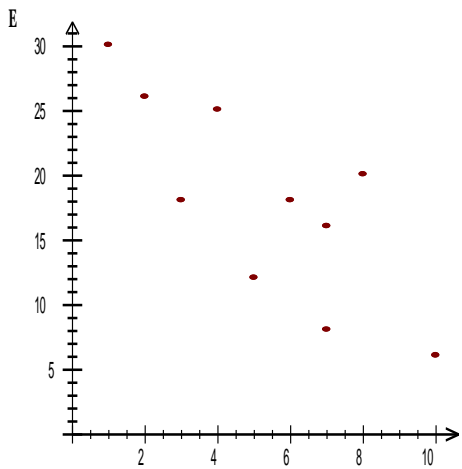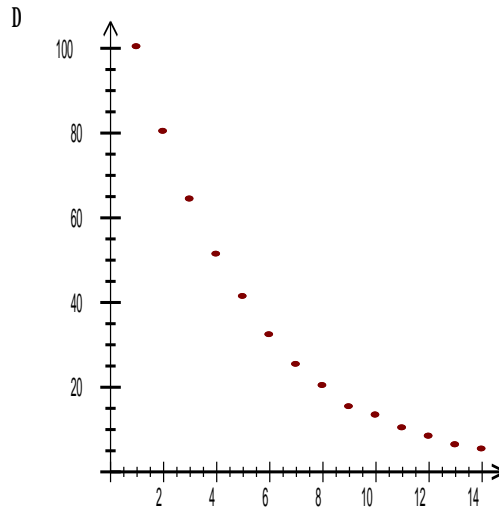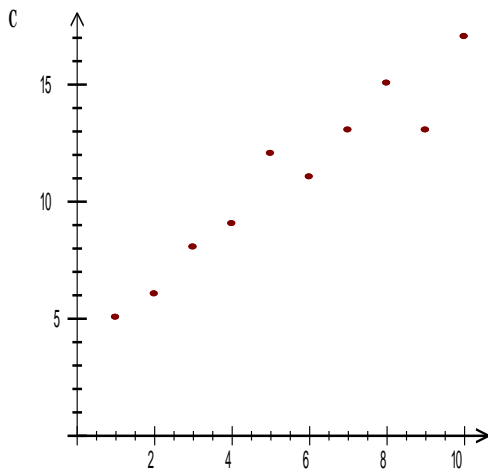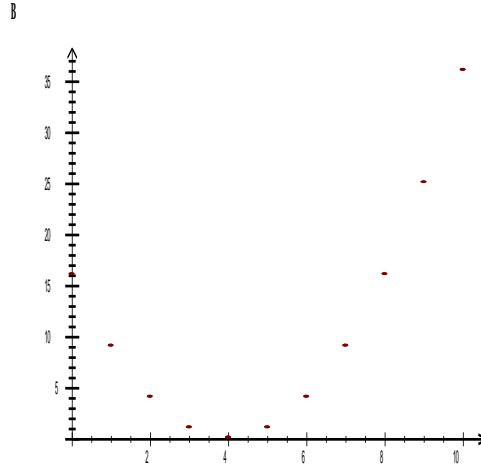
Order graphs A, B, C, D and E in the box below.

b)   Which two graphs will have the higest coefficent of determination?

Question Six: [1, 2, 3: 6 marks]

1 000 children across the state were asked to complete an appitute test and an IQ test produced by a particular company. A random sample of 10 student's appitude test score and their corresponding IQ test score are shown in the table below.

Student	Apptitude test Score	IQ test score
1	12	14.4
2	99.7	122.4
3	60.9	57
4	53.2	33.6
5	65.3	51.6
6	49.6	42.8
7	20.4	-5.4
8	26.6	18.5
9	23.2	46.7
10	19	39.8

a) Calculate the correlation coefficient, $r_{IQ\,AT}$.

b) Describe the association between the apptitude test score and the IQ test score with reference to the correlation coefficient.

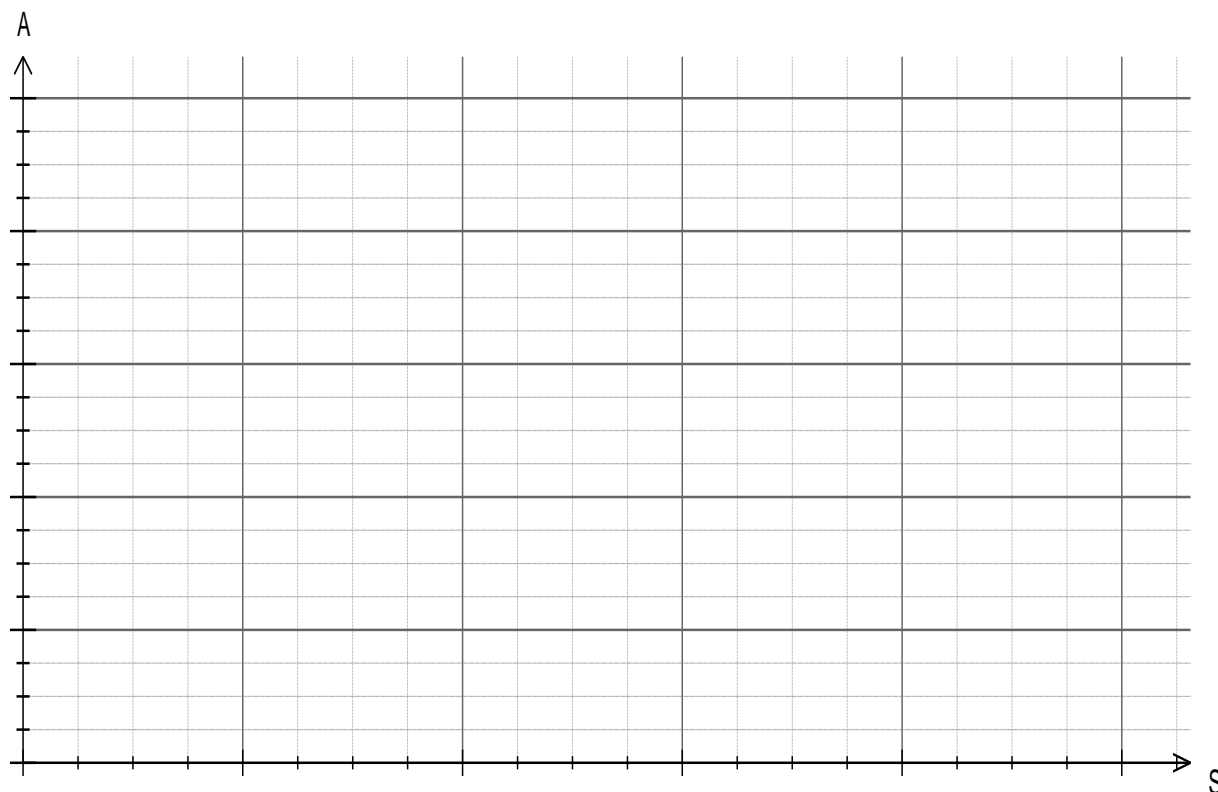c) Calculate the coefficient of determination and describe what this value means.

Question Seven: [3, 1, 1, 2, 1, 3, 3: 14 marks]

The following figures have been obtained from School Curriculum and Standard Authority (SCSA) website. The data shows the final school examination result and corresponding ATAR examination result for 13 random students.

Student	Final School Exam % (S)	ATAR Exam % (A)
A	56	62
B	74	69
C	63	66
D	23	29
E	61	70
F	60	68
G	59	62
H	77	79
I	81	80
J	49	51
K	70	73
L	66	70
M	77	80

a)   Draw a scatterplot for the data above.

b)   Identify the response variable.

c)   Calculate the correlation coefficient, r_{SA}.

d)   Describe the association between the final school exam score and the ATAR score
        with reference to your answer in part c).

e)   Calculate the coefficient of determination.

f)   Describe what the value calculated in part e) means.

More data is found showing the same 13 student's final school marks, for the same subject,
as a percentage. This data is shown in the following table.

Student	Final School Mark %
A	60
B	73
C	65
D	30
E	69
F	66
G	60
H	80
I	80
J	50
K	74
L	69
M	78

g)   Does this score provide a closer match to predicting the student's ATAR scores?
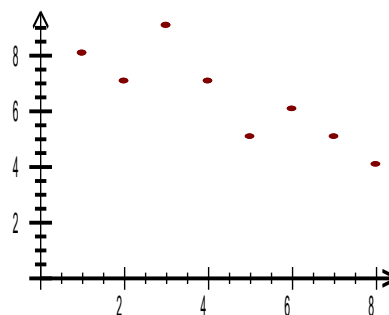   (Justify your answer with relevant statistics)

education equals

Question One: [4 marks]

Match the correlation coefficient to the correct scatter plot.

Correlation coefficient (r)	−0.84	0.40	0.95	−0.18
Scatter plot	B ✔	C ✔	A ✔	D ✔

Question Two: [2, 2, 2, 2: 8 marks]

Describe the likely relationship between each of pair of variables.

a)   Height vs the weight maths Applications students in your class.

Strong, positive, linear relationship

✓ ✓

b)   Value of a 2015 Holden Commadore vs time.

Strong, negative, linear relationship

✓ ✓

c)   Fuel consumption of a vehicle vs the kilometers it has travelled.

Strong, positive, linear relationship

✓ ✓

d)   Hours spent studying for maths Applications vs the test results obtained.

Moderately strong, positive, linear relationship

✓ ✓

Question Three: [1, 1, 1: 3 marks]

Match the likely correlation coefficient, in the table at the bottom of the page, with the two variables described.

a)   Alcohol consumption vs liver function

$r = -0.89$ ✓

b)   Birth rate (number of births per 1000 population) vs female life expectancy

$r = -0.70$ ✓

c)   Energy (Kj) vs fat content (g)

$r = 0.90$ ✓

$r = -0.89$	$r = 0.56$	$r = 0.90$	$r = -0.70$

Question Four: [2, 3: 5 marks]

Some people feel that it is more difficult to sleep when temperatures are higher. It has been claimed that, "High overnight temperatures cause people to sleep less than the standard 10 hours."

a)   State the explanatory and the response variable in the above claim.

Explanatory – temperature   ✓

Response – hours slept   ✓

The table below shows the hours slept and the maximum temperature reached overnight for a sample of 18 people.

| Max Temperature Over Night	Number of Hours slept
10	9
14	8
11	9
12	9
9	13
20	6
15	7
15	6
16	8
17	6
8	12
11	7
21	6
19	5
18	6
18	7
13	7

b)   Given the above data, comment on the claim made. Consider the use of statistical measures to support your answer where necessary.

$r = -0.82 \, (2dp)$ ✓   ✓

The correlation coeffient is strong, indicating that there is a strong relationship between the variables but we cannot conclude that higher temperatures will result in less sleep. This data indicates a relationship but cannot point to causality. ✓

Question Five: [3, 2: 5 marks]

a)   Order the following scatter graphs from strongest positive linear relationship to
      weakest negative linear relationship.

B

A

C

D

E
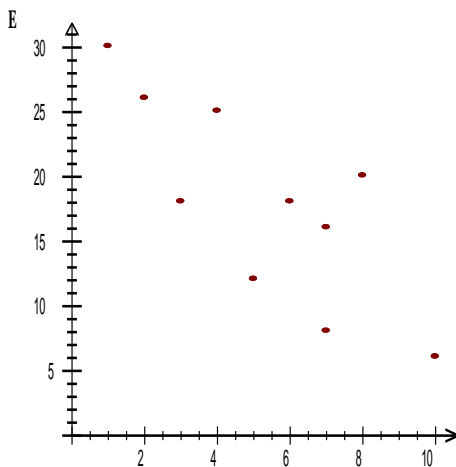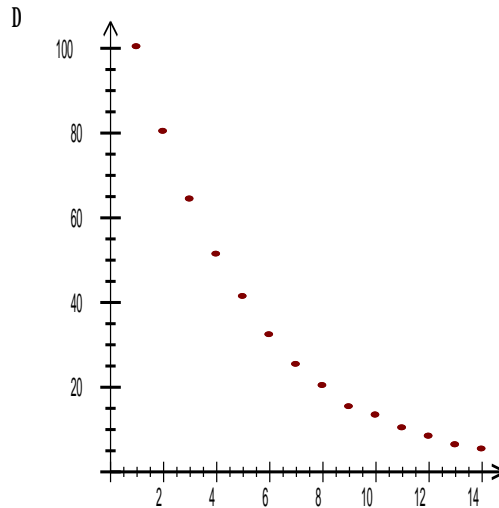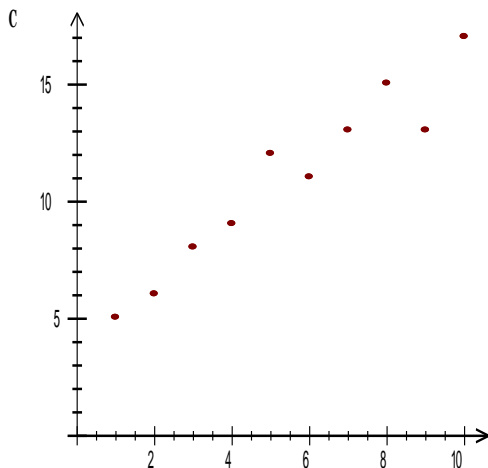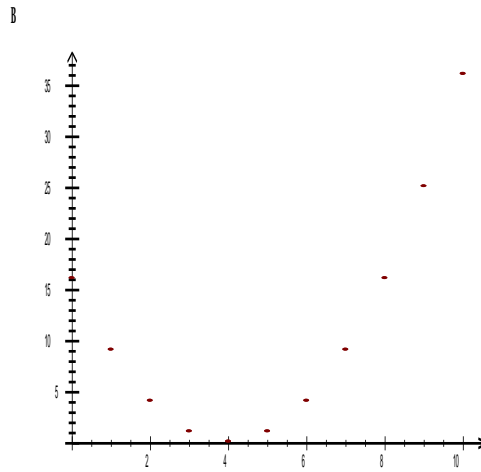
Order graphs A, B, C, D and E in the box below.

C, A, B, E, D
✔✔✔

b)   Which two graphs will have the higest coefficent of determination?
      C and D   ✔ ✔

Question Six: [1, 2, 3: 6 marks]

1 000 children across the state were asked to complete an appitute test and an IQ test produced by a particular company. A random sample of 10 student's appitude test score and their corresponding IQ test score are shown in the table below.

Student	Apptitude test Score	IQ test score
1	12	14.4
2	99.7	122.4
3	60.9	57
4	53.2	33.6
5	65.3	51.6
6	49.6	42.8
7	20.4	-5.4
8	26.6	18.5
9	23.2	46.7
10	19	39.8

a)   Calculate the correlation coefficient, $r_{IQ\ AT}$.

$$r = 0.8529\ (4\ dp) ✔$$

b)   Describe the association between the apptitude test score and the IQ test score with reference to the correlation coefficient.

There is a strong, positive, linear relationship ✔ ✔

c)   Calculate the coefficient of determination and describe what this value means.

$r^2 = 0.7274$ ✔

Approximately 73% of variation in the appititde test scores can be attributed to the variation in the IQ test scores, the other 27% is unexplained. ✔ ✔

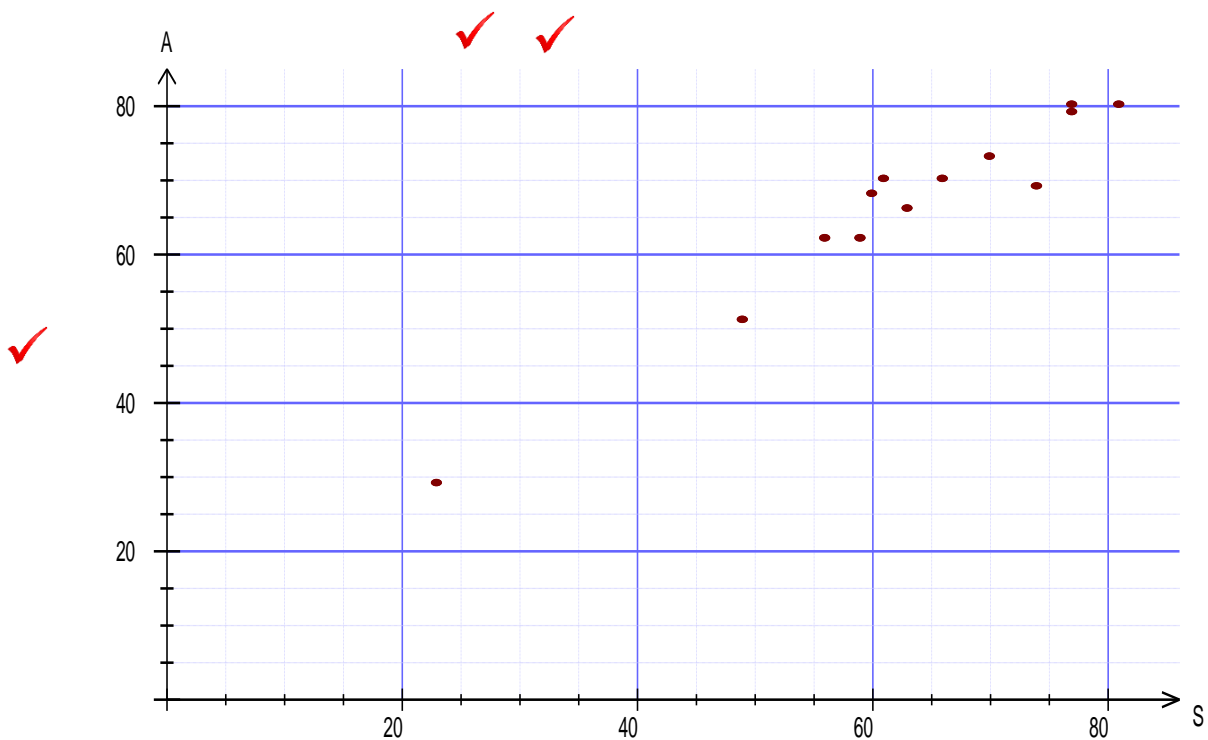**Question Seven: [3, 1, 1, 2, 1, 3, 3: 14 marks]**

The following figures have been obtained from School Curriculum and Standard Authority (SCSA) website. The data shows the final school examination result and corresponding ATAR examination result for 13 random students.

Student	Final School Exam % (S)	ATAR Exam % (A)
A	56	62
B	74	69
C	63	66
D	23	29
E	61	70
F	60	68
G	59	62
H	77	79
I	81	80
J	49	51
K	70	73
L	66	70
M	77	80

a)   Draw a scatterplot for the data above.

b)   Identify the response variable.

ATAR Exam %

c)   Calculate the correlation coefficient, r_{SA}.

$$r = 0.9727 \ (4dp)$$ ✔

d)   Describe the association between the final school exam score and the ATAR score
        with reference to your answer in part c).

✔ ✔

There is a very strong, positive, linear relationship between the Final Exam score and   the
ATAR exam score.

e)   Calculate the coefficient of determination.

$$r^2 = 0.9461$$ ✔

f)   Describe what the value calculated in part e) means.

✔ ✔

Approximately 95% of the variation in the ATAR scores between these students can be
explained by the variation in the final school exam scores. The other 5% is unexplained. ✔

More data is found showing the same 13 student's final school mark, for the same subject, as
a percentage. This data is shown in the following table.

Student	Final School Mark %
A	60
B	73
C	65
D	30
E	69
F	66
G	60
H	80
I	80
J	50
K	74
L	69
M	78

g)   Does this score provide a closer match to predicting the student's ATAR scores?
        (Justify your answer with relevant statistics)

✔ ✔ $r = 0.9919 \ (4 \ dp)$ Since this is higher than the correlation coefficient between the final
school exam % and the ATAR %, this does provide a more reliable prediction. ✔