Student Name: Marking Key

Eastern Goldfields College

Mathematics Applications U3&4 2017

Investigation 1 Validation

Working Time: 50 minutes **CALCULATORS ARE ALLOWED** Total Marks: 24

Bivariate Data Analysis

Question 1 (9 marks: 2, 2, 5)
The table below shows the hours spent by a sample of 20 high school students watching television and doing homework in a week.

Hours watching TV	4	5	5	0	9	6	20	4
Hours of homework	10	5	1	2	3	7	9	4
Hours watching TV	6	2	40	5	1	8	17	2
Hours of homework	4	1	1	1	1	10	20	1
Hours watching TV	13	1	14	14				
Hours of homework	14	6	0	9				

Some people feel that TV is a distraction to young high school students; in fact, they would claim, 'the more hours spent by a student watching television, the less hours the student will spend doing homework'.

Complete the following questions and use these data to comment on the above statement.

(a) Decide which is the explanatory variable and which is the response variable in the above comment. (2 marks)

E: TV Watching Hours ✓

R: Homework Completion Hours ✓

(b) Calculate the statistical measures you will use to decide whether the proposition is reasonable or not. (2 marks)

$r = 0.1983815$ ✓

$r^2 = 0.0393552$ ✓

$\simeq 3.9\%$

(c) Use your measures from (b) to comment on the proposition, 'the more hours spent by a student watching television, the less hours the student will spend doing homework'. (5 marks)

$r = 0.1984 \Rightarrow$ weak, positive linear relationship.

not reasonable
not related
not true

$r^2 = 0.0394 \simeq 3.9\% \Rightarrow$ Hours watching TV has a small effect on HW completed

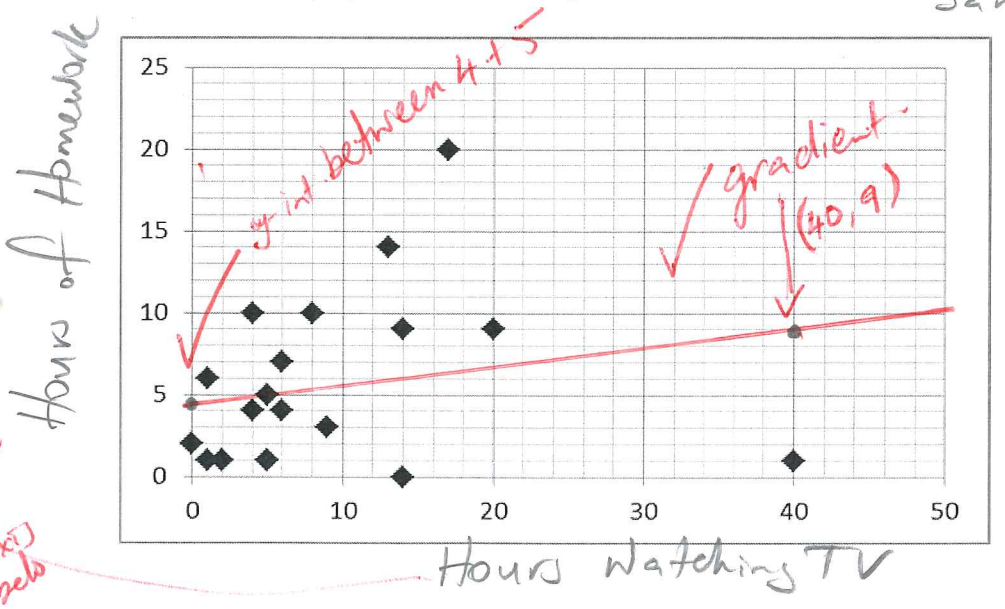
sample of data does not support the proposition

proposition $\Rightarrow$ a negative relationship. As hours of TV watching increases, Homework hours decreases.

sample not support this.

Question 2 (10 marks: 3, 4, 3)

The diagram below is the scatter plot for the sample in the table above.



(a) Label the axes and draw the regression line from 1(b) on the graph. (3 marks)

(b) Calculate the equation of the linear regression line, using the variables defined here: Hours watching TV = t and Hours of Homework = w (4 marks)

$w = 0.1124t + 4.4609$ (4 dp)

$w = 0.11t + 4.46$

(c) Using the line of best fit in (b), calculate the residual for the student who watched 17 hours of TV and spent 20 hours doing homework. (3 marks)

$\hat{w} = 0.11(17) + 4.46$

$\simeq 6.33$ (2dp) F.T.

Residual $\simeq 20 - 6.33 \simeq 13.67$ hours

Question 3 (5 marks: 1, 1, 1, 1, 1)

Below are the four stages of the Investigation Process. Use the information and data within this validation to address each of these stages.

Stage 1: *Clarify the problem and formulae one or more questions that can be answered with data.*

a) Is this validation, what was the problem and questions you were required to answer?

½ problem or proposition = The more hours spent by a student watching television, the less hours the student will spend studying

½ Question = Does hours a student spends watching TV explain the number of hours a student studies?

OR Is there a relationship?

Stage 2: *Design and implement a plan to collect and obtain appropriate data.*

b) Data for this validation was given to you. How could you have collected this data if it was not provided?

survey ✓ a wide variety of students from different ages, schools,

Stage 3: *Select and apply appropriate graphical or numerical techniques to analyse the data.*

c) In this validation, what graphical and numerical techniques were used to analyse the data?

scattergraph
correlation coefficient (r)
coefficient of determination (r^2)] OR ??
linear regression line
residual.

✓ = lists 4-5
½ = lists 2-3
0 = lists 0-1

Stage 4: *Interpret the results of this analysis and relate the interpretation to the original question and communicate findings in a systematic and concise manner.*

d) In this validation, what are the findings?

✓ no relationship between the hours a student watches TV and hours the student spends studying.

e) What could be done to improve your analysis and findings?

✓ Lists any 1 reasonable answer

eg:

- larger sample

OR

- wider sample eg. tertiary (uni/TAFE) st's.

OR

- etc

TAKE HOME

0 marks – Not Submitted

1 mark – attempted with lots of / mostly errors / incomplete

2 marks – in between ↑↓ – minor errors / mostly completed.

3 marks –
- table of downloaded data
- ⊕ scattergraph of ↑
- ⊕ Explan + Response variables
- ⊕ Outliers identified + removed (if any)
- ⊕ r, r^2, $y = mx + c$,
- ⊕ table of class data
- ⊕ predictions + residuals
- ⊕ interpretation.

~ END OF VALIDATION ~