TRINITY COLLEGE

Year 12 Mathematics Applications
Test 1 2018

Calculator Assumed

Bivariate Data

STUDENT'S NAME _____

DATE: Thursday 1st March **TIME:** 45 minutes **MARKS**: 45

INSTRUCTIONS:

Standard Items: Pens, pencils, drawing templates, eraser

Special Items: Three calculators, notes on one side of a single A4 page (these notes to be handed in with this assessment)

Questions or parts of questions worth more than 2 marks require working to be shown to receive full marks.

---

1. (3 marks)

A scatterplot is drawn for the relationship between variables x and y.

(a) Comment on the relationship between x and y. [2]

(b) Give a practical example for the relationship which is modelled in the scatterplot above. [1]

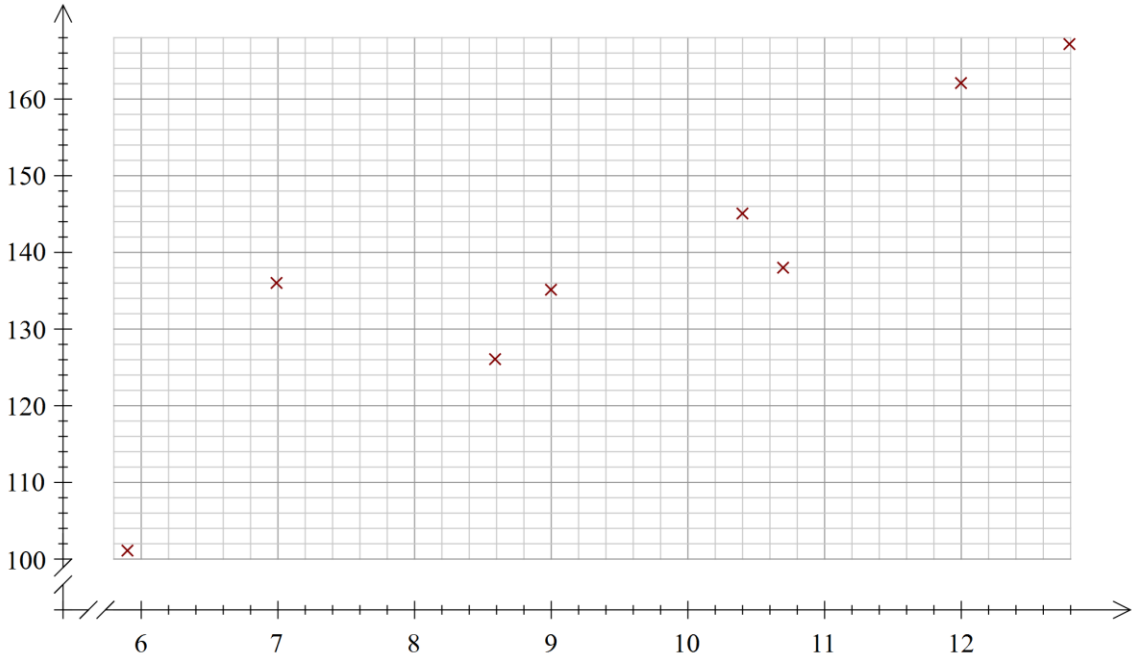
2. (21 marks)

A Fantasy Football participant decided to determine whether there was a relationship between the transfer cost, in millions of pounds, of a player and the total points they scored for the season. He collected the following data:

Player Cost	10.4	12	9	8.6	5.9	10.7	12.8	7	**9.2**	**11.3**	**12**
Points Scored	145	162	135	126	101	138	167	136	**139**	**140**	**162**

(a) State the explanatory variable. [1]

(b) Complete the scatterplot below by plotting the **last 3 data points** and labelling the horizontal axis and the vertical axis clearly. [3]



(c) Calculate the correlation coefficient, r and comment briefly on the relationship between the variables. [4]

(d) (i) Determine the equation for the least-squares line that models these data. State the slope and vertical intercept. [2]

(ii) Draw the line from (d)(i) on the scatterplot in part (b) indicating two calculated points on the graph. [2]

(iii) Interpret the slope of the least-squares line. [2]

(e) Calculate the coefficient of determination and interpret it. [2]

(f) Estimate the total points a player costing 8 million pounds would score. Comment on the reliability of this estimate. [3]

(g) The participant concluded that because of the strength of the correlation, "the higher a player's transfer cost the more points that player will score". Comment on the appropriateness of this statement. [2]

3. (4 marks)

A study into the effectiveness of a pain relief drug was conducted. The response time (R minutes) was measured for different doses of the drug (D milligrams). The results of the study were used to obtain the least squares regression line $R = -8.54D + 45.55$, and the coefficient of determination was found to be 0.961.

(a) Will the correlation coefficient between D and R be positive or negative? Give a reason for your answer. [2]

(b) Determine the correlation coefficient between D and R. [2]

4. (8 marks)

Five hundred University students were classified according to their personality (introvert or extrovert) and were then asked about their hobby preference (reading, gardening, exercise or watching TV).

	Hobby				
Personality	Reading	Gardening	Exercise	Watch TV	TOTAL
Introvert	42	50		86	
Extrovert			45		300
TOTAL	231			121	

(a) Complete the missing entries in the table. [2]

(b) Which variable is the explanatory variable? [1]

(c) Construct a percentage two-way table that will help identify whether hobby preference is associated with personality. **Round values to the nearest percentage.** [3]

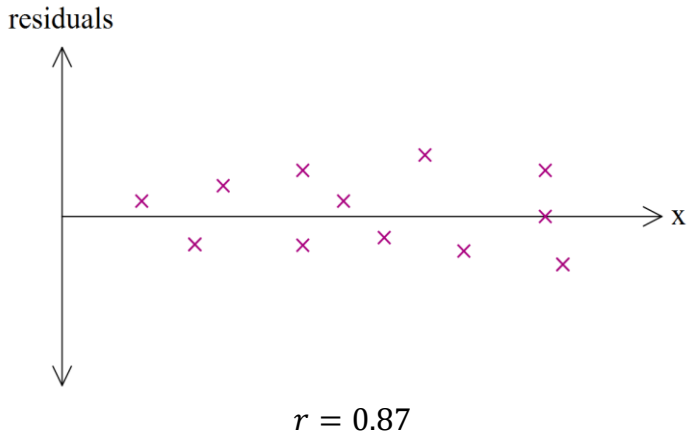
	Hobby			
Personality	Reading	Gardening	Exercise	Watch TV
Introvert				
Extrovert				

(d) Is there an association between the response and explanatory variables? Provide evidence to support your decision. [2]

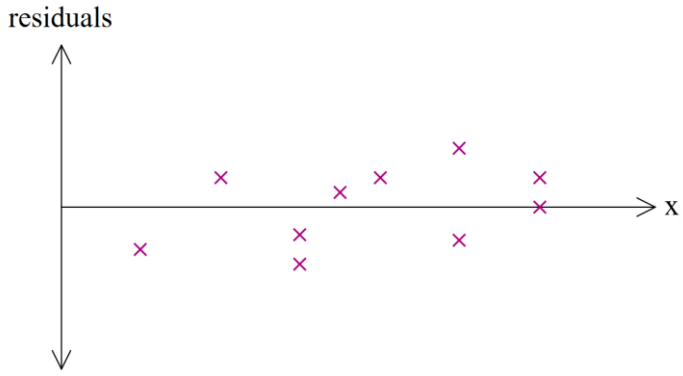
5. (10 marks)

(a) Three different sets of data were collected. For each set, the line of regression, correlation coefficient and residuals were determined. Shown below are the residual plots for each set of data, along with the correlation coefficient.

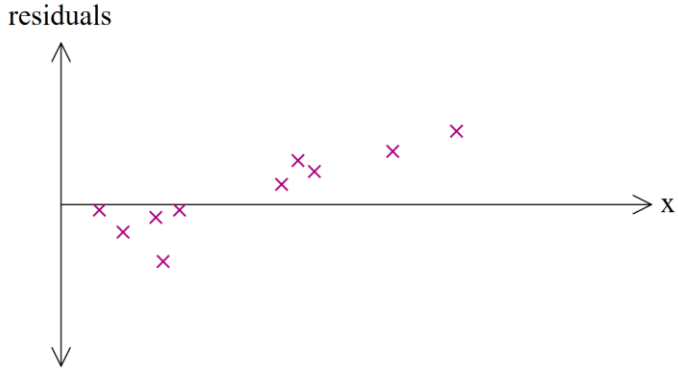
In each case, state (with reasoning) whether the linear regression model used would be suitable for prediction purposes. [6]

<table>
<tr>
<td> 

$r = 0.87$</td>
<td>Is the model appropriate?

Reason</td>
</tr>
<tr>
<td> 

$r = -0.23$</td>
<td>Is the model appropriate?

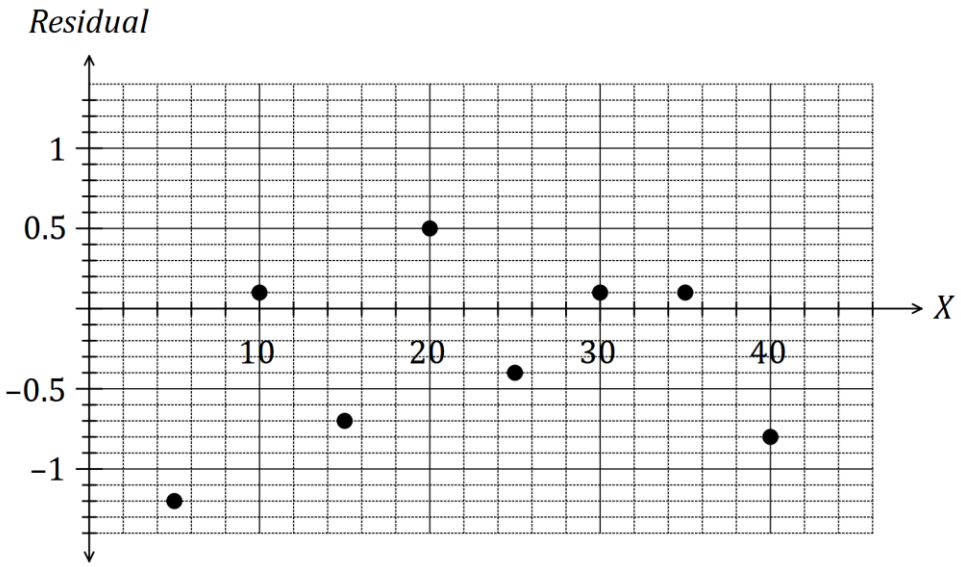
Reason</td>
</tr>
<tr>
<td> 

$r = 0.91$</td>
<td>Is the model appropriate?

Reason</td>
</tr>
</table>

(b) A linear model was fit to a set of data (X, Y) and the equation of the least squares regression line was found to be $\hat{Y} = 1.25X + 20.72$.

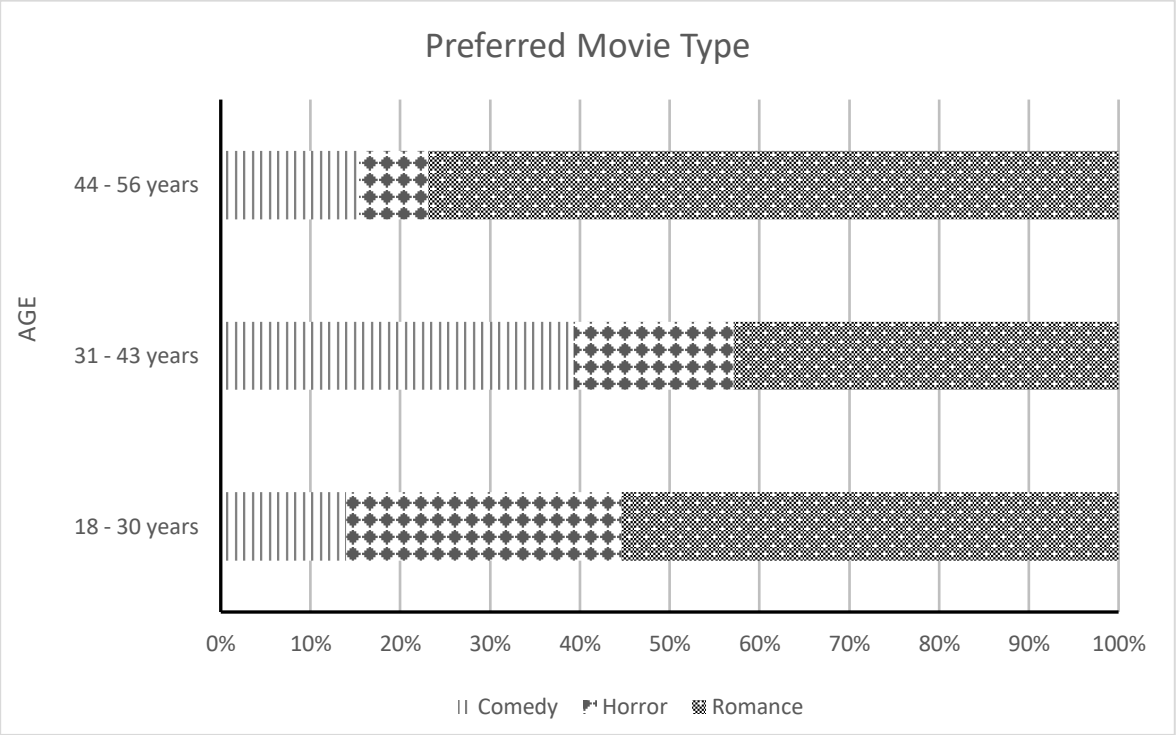
A plot of the residuals for this model is shown on the axes below.



Use the least squares regression line and the residual plot to determine the **actual** value of Y when $X = 15$. [3]

6. (3 marks)

The proportional bar graph below shows the preferred movie type for a group of people aged between 18 and 56 years.



Discuss whether any association appears to exist between the preferred movie type and age of person surveyed. Justify your decision.