

|  |  |  |
| --- | --- | --- |
| Mathematics Department | |  |
| Course: ATMAA | |
| Topic Title: Bivariate data  Test 2 | |
| Student Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | Date: \_\_\_\_\_\_\_\_\_\_\_\_ | | |
| Special Instructions: **Calculator Allowed**  1 page of A4 notes and Formula Sheet Allowed | Time Allowed: 60 mins | | |
|  | Marks: / 58 | | |
|  |  | | |

**Question 1.** (3, 2, 2, 5, 2: 14 marks)

315 students were interviewed about their favourite subject – Mathematics or English. Of the 170 girls, 70 preferred Mathematics, while 45 boys preferred English.

1. Construct a two way table showing this information.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. Determine the explanatory and response variable.
2. Construct an appropriate percentage two way table.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. Construct a segmented column graph.

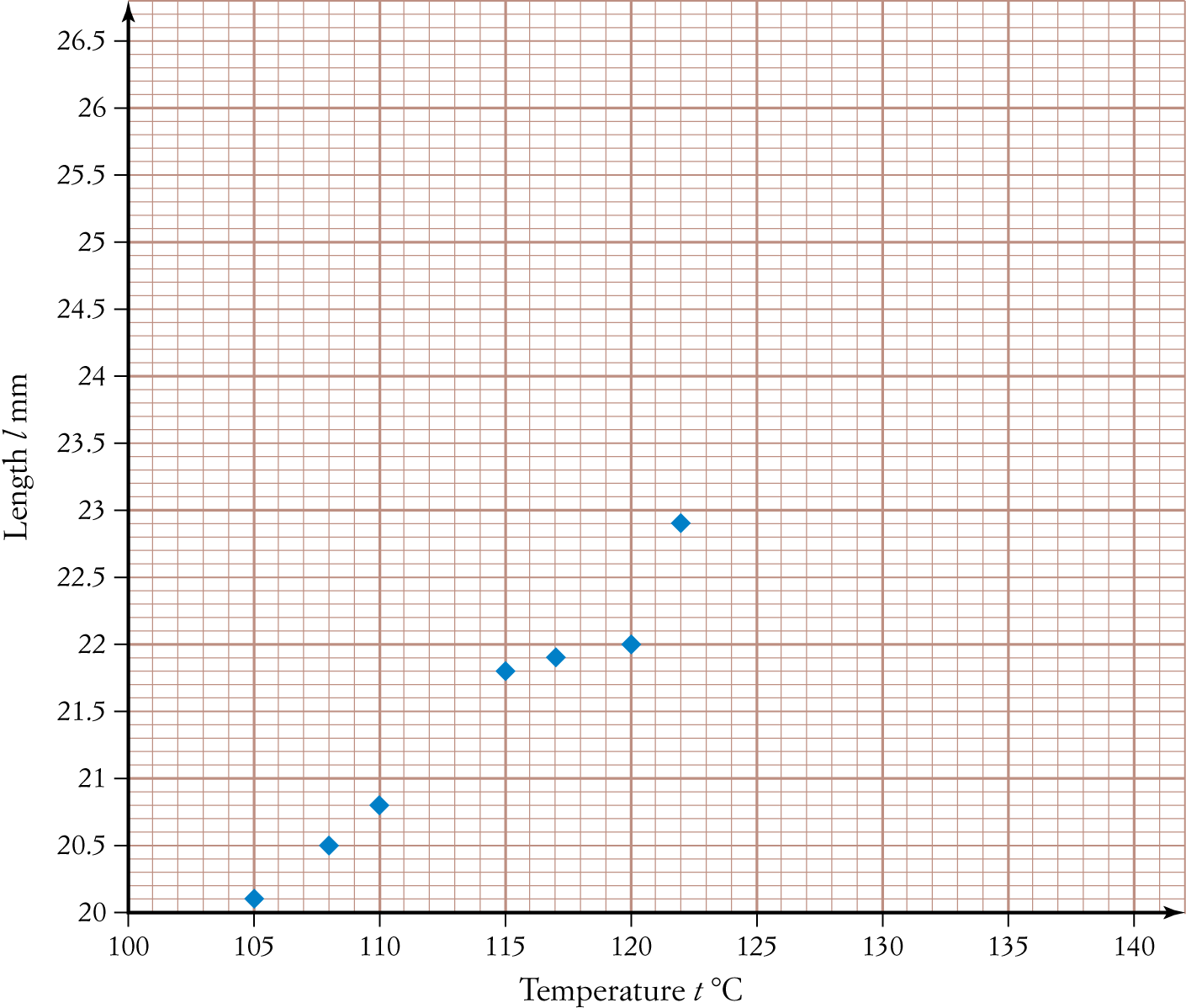


1. Is there an association between the two variables? Explain your answer.

**Question 2.** (1, 1, 2: 4 marks)

The length (*l* mm) of a metal bar was measured at various temperatures (*t*°C) to give the following results.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **t°C** | 105 | 108 | 110 | 115 | 117 | 120 | 122 | **125** | **130** | **135** |
| ***l* mm** | 20.1 | 20.5 | 20.8 | 21.8 | 21.9 | 22.0 | 22.9 | **24.3** | **25.1** | **26.1** |

a) Complete the following scatter plot by adding the last three points shown in bold in the table.

b) Calculate Pearson’s correlation coefficient for the data, correct to two decimal places.

c) Interpret the relationship between the variables ‘Length’ and ‘Temperature’, referring to form, direction and strength.

**Question 3.** (3, 2, 2, 2: 9 marks)

Data was collected to investigate whether a person’s income ($) depends on their height (cm) and is displayed in the table below. State all answers correct to two decimal places.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Height (cm)** | 174 | 162 | 165 | 183 | 185 | 158 | 176 | 196 |
| **Income ($)** | 62 000 | 36 000 | 44 000 | 51 000 | 60 000 | 47 000 | 39 000 | 26 000 |

a) Using CAS, calculate the least-squares regression line that models the data.

b) Calculate and interpret Pearson’s correlation coefficient for the data.

c) Calculate and interpret the coefficient of determination for the data.

d) Is the regression line found in part **a** an appropriate linear model? Justify your answer.

The following information relates to Questions 4 and 5.

A group of 25 year olds were surveyed regarding the number of hours spent exercising per week and

their resting heart rate (beats per minute). The data collected for the number of hours of exercise

per week ranged from 0 hours to 6 hours.

The regression line that models the data was found to be:

Resting heart rate = 63 − 0.96 × Exercise,

with a correlation coefficient of −0.52.

**Question 4.** (7 marks)

Use the information above to complete the following sentences.

a) A person’s resting heart rate will be \_\_\_\_\_\_\_\_\_\_ beats per minute when the amount of

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ is 0 hours per week.

b) \_\_\_\_\_\_\_\_\_\_\_ of the variation in \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ can be explained by the

variation in the amount of \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

c) A person’s resting heart rate will \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ by \_\_\_\_\_\_\_\_\_\_\_beats per minute

for every one hour increase in exercise.

**Question 5.** (2, 2, 2: 6 marks)

Use the regression line to:

a) Predict the resting heart rate for a person who exercises for 3 hours per week. Is this prediction classified as interpolation or extrapolation? State answer correct to one decimal place.

b) Predict the amount of exercise required per week for a person with a resting heart rate of 55 beats per minute. Is this prediction classified as interpolation or extrapolation? State answer correct to one decimal place.

c) Which of the above predictions is the most reliable? Justify your answer.

**Question 6.** (3, 2, 2: 7 marks)

The weights (kg) of 13-year-old students (*y*) and the number of chocolate bars consumed per week (*x*) were recorded and the values of the following statistics were determined.

= 3.23 = 1.6 = 41.55 = 5.61 *r* = 0.56

a) Calculate the least-squares regression line that models these data, stating all coefficients correct to two decimal places.

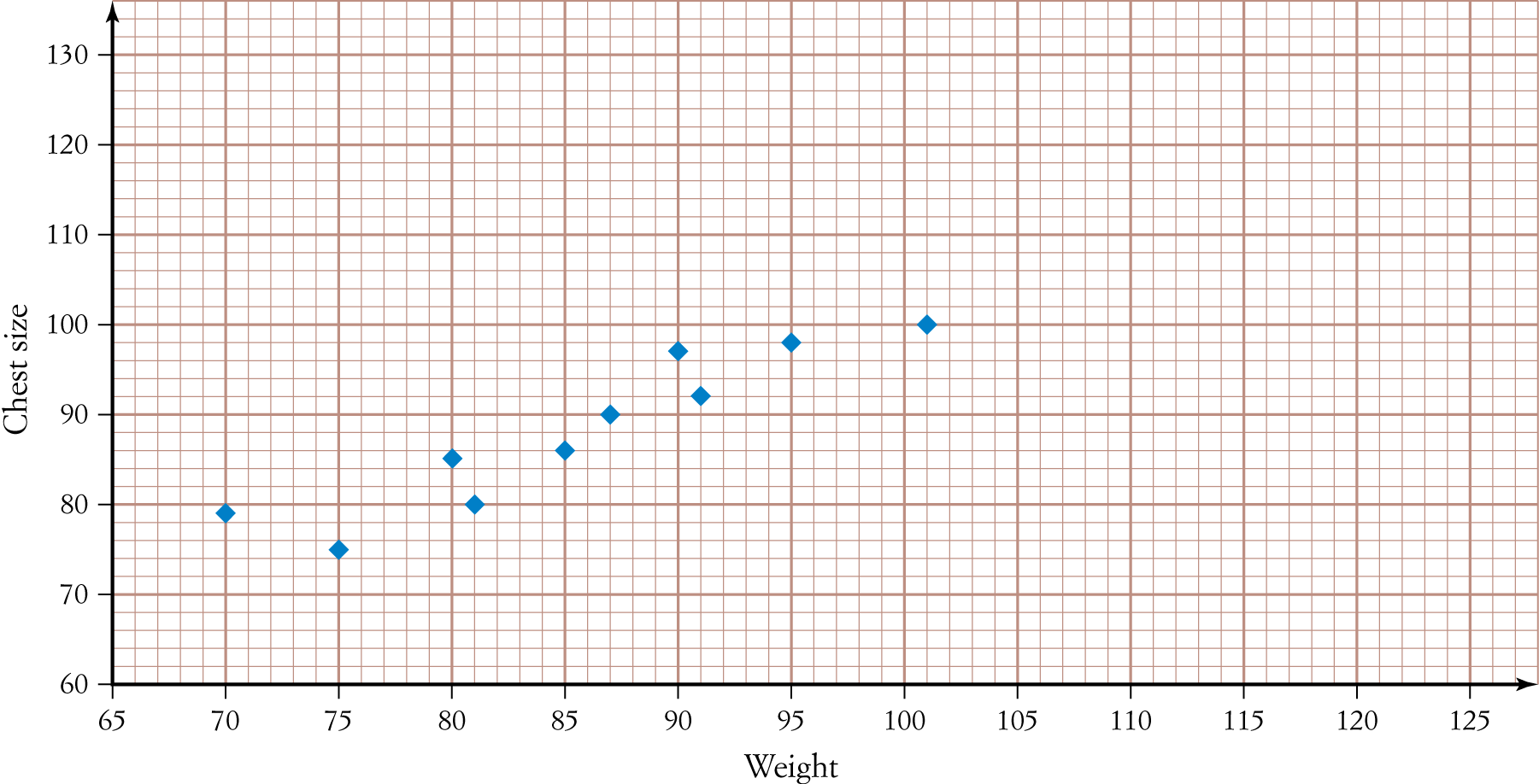
b) Use the regression line to predict the weight of a 13-year-old student who eats 5 chocolate bars per week, correct to two decimal places.

c) Hence, calculate the residual value for Peter who weighs 43 kg and eats 5 chocolate bars per week.

**Question 7.** (2, 2, 2, 2, 2, 1: 11 marks)

Data was collected from a rugby team to investigate whether there is a relationship between a player’s weight (kg) and their chest size (cm). The results are displayed in the table below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Weight** | 70 | 75 | 80 | 81 | 85 | 87 | 90 | 91 | 95 | 101 | **102** | **105** | **115** | **120** |
| **Chest size** | 79 | 75 | 85 | 80 | 86 | 90 | 97 | 92 | 98 | 100 | **99** | **103** | **102** | **123** |

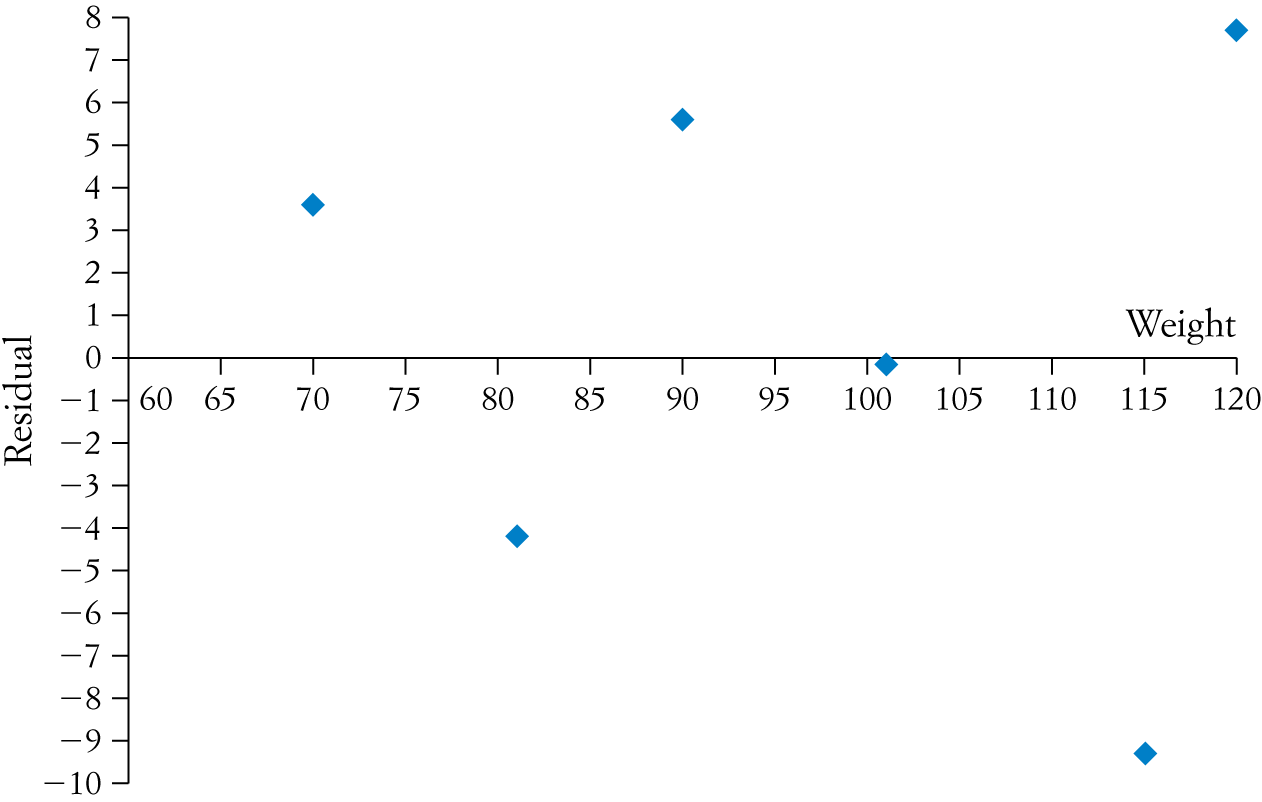
a) Add the last four data points from the table (**in bold**) to the scatter plot below.

b) Calculate the least-squares regression line that models the data, stating all coefficients correct to two decimal places.

c) Hence, sketch the regression line onto the scatter plot in part **a**.

d) Calculate all the residual values (to one decimal place) for the data and complete the table below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Weight** | 70 | 75 | 80 | 81 | 85 | 87 | 90 | 91 | 95 | 101 | 102 | 105 | 115 | 120 |
| **Residual** | 3.6 |  |  | -4.2 |  |  | 5.6 |  |  | -0.2 |  |  | -9.3 | 7.7 |

1. Hence, complete the residual plot below.
2. Using the residual plot, justify whether the regression model found in part **b** is appropriate.