# TRINITY COLLEGE

## Year 12 Mathematics Applications
## Test 1     2021
### Calculator Assumed
### Bivariate Data

STUDENT'S NAME _MARKING KEY_

**DATE:** Friday 26th February          **TIME:** 50 minutes          **MARKS:** 48

**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.      (5 marks)

The relationship between a data collection of weight ($w$) and units of horsepower ($h$) of a car, with weights ranging between 1300 kg and 2000 kg obtained the regression equation $h = -10.5 + 0.09w$.

(a)     Predict the horsepower of a car that weighs 2100 kg.                    [1]

$178.5$          correct ans ✓

(b)     A prediction was made which resulted in a horsepower of 151.5 units. What weight was the prediction based on?                    [2]

$151.5 = -10.5 + 0.09w$          substitute into equation ✓

$w = 1800$ kg.          correct answer ✓

(c)     Consider which of the three predictions in parts (a) and (b) would be least reliable? Justify your answer.                    [2]

(a) is least reliable due to extrapolation

✓ (a)          ✓ extrapolation mentioned.

2.      (6 marks)

The data in the table below shows the country of origin of cars parked at three different shopping centres on a weekday morning. The shopping centres A, B and C are located respectively in high, middle, and low socioeconomic suburbs.

|            | A   | B   | C   |
|------------|-----|-----|-----|
| Australian | 75  | 62  | 39  |
| German     | 157 | 78  | 25  |
| Korean     | 46  | 126 | 117 |
| Japanese   | 85  | 143 | 108 |
| Other      | 64  | 68  | 54  |

(a)     Is the socioeconomic class of the suburb the explanatory or response variable?     [1]

Explanatory                                    ✓ correct ans

(b)     Complete the appropriate row/column percentage table to represent this data, to the nearest whole number.     [3]

|            | A    | B    | C    |
|------------|------|------|------|
| Australian | 18%  | 13%  | 11%  |
| German     | 37%  | 16%  | 7%   |
| Korean     | 11%  | 26%  | 34%  |
| Japanese   | 20%  | 30%  | 31%  |
| Other      | 15%  | 14%  | 16%  |

✓ column %.

✓ correct %.

✓ rounded.

(c)     State, with justification, if there is an association between the cars' country of origin and the suburb of the shopping centre.     [2]

✓ states there is an association

✓ Justifies with reference to %.

e.g. "37% of the cars in suburb A are from Germany as opposed to 16% of the cars from suburb B and 7% from suburb C."

3.    (8 marks)

Mr Presser analysed the relationship between the scores of ten Chemistry students from 2020 in their Mock exam ($m$) in September and then their ATAR ($a$) exam in November, establishing the data to have a correlation coefficient of 0.9794 and a least squares regression equation of $\hat{a} = 1.0589m + 0.45$.

The table below shows the values of the Mock exam ($m$), the ATAR exam ($a$), the predicted scores ($\hat{a}$) and the residuals ($a - \hat{a}$).

| $m$ | 38 | 41 | 45 | 49 | 54 | 57 | 60 | 63 | *D* | 73 |
|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | 41 | 44 | *C* | 52 | 56 | 60 | 65 | 61 | 78 | 80 |
| $\hat{a}$ | 40.692 | 43.869 | 48.105 | 52.341 | 57.635 | *A* | 63.989 | 67.165 | 75.637 | 77.755 |
| $a - \hat{a}$ | 0.308 | 0.131 | 2.895 | -0.341 | -1.635 | *B* | 1.011 | -6.165 | 2.363 | 2.245 |

(a)    For the table, calculate the values for *A, B, C* and *D*.    [4]

√√√√
one mark
per value.
=T with
working.

$A = 1.0589(57) + 0.45$

$A = 60.807$

$B = 60 - 60.807$

$B = -0.807$

$2.895 = C - 48.105$
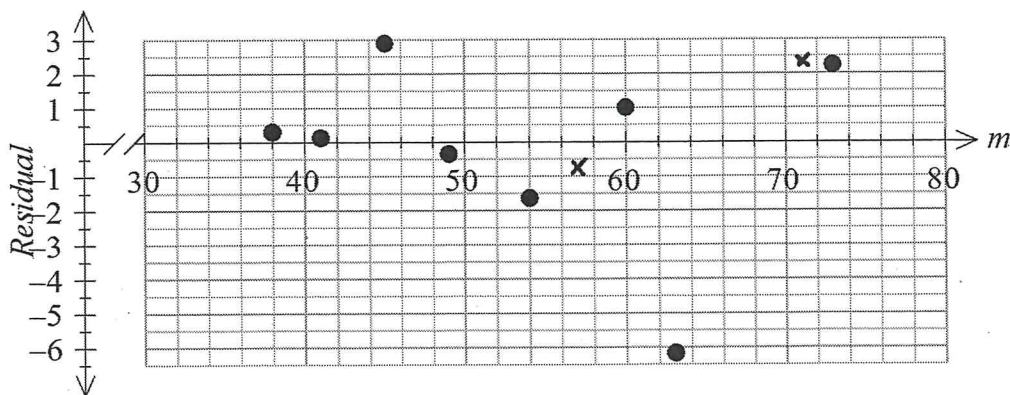
$C = 51$

$75.637 = 1.0589D + 0.45$

$D = 71$

(b)    Add the missing two points to the residual plot on the grid provided.    [1]



(c)    Does the residual plot indicate that linear regression is appropriate? Explain your answer.    [2]

Yes, due to the random nature of the plot.    √ Yes    √ Explanation.

(d)    The residual of the student who scored 63 in the Mock exam and 61 in the ATAR exam shows the point to be significantly further away from the least squares regression line than any other point. If this point was removed would the gradient of the least-squares line increase, decrease or remain the same?    [1]

increase    √ correct ans.

4.    (7 marks)

The table below shows the number of times each PSA school has won each of the summer sports between 1980 and 2019.

|  |  | Basketball | Cricket | Rowing | Tennis | Volleyball | Water Polo | Total |
|---|---|---|---|---|---|---|---|---|
| Aquinas | Number of Wins | 13 | 7 | 16 | 7 | 16 | 13 | 71 |
|  | Percentage of Wins | 18.3% | 9.9% | 22.5% | 9.9% | 22.5% | 18.3% | 100% |
| Christchurch | Number of Wins | 3 | 2 | 2 | 12 | 0 | 5 | 24 |
|  | Percentage of Wins | 12.5% | 8.3% | 8.3% | 50% | 0% | 20.8% | 100% |
| Guildford | Number of Wins | 4 | 3 | 3 | 1 | 0 | 1 | 12 |
|  | Percentage of Wins | 33.3% | 25% | 25% | 8.3% | 0% | 8.3% | 100% |
| Hale | Number of Wins | 10 | 9 | 2 | 10 | 1 | 4 | 36 |
|  | Percentage of Wins | 27.7% | 25% | 5.6% | 27.7% | 2.7% | 11.1% | 100% |
| Scotch | Number of Wins | 5 | 3 | 5 | 8 | 0 | 5 | 26 |
|  | Percentage of Wins | 19.2% | 11.5% | 19.2% | 30.8% | 0% | 19.2% | 100% |
| Trinity | Number of Wins | 8 | 4 | 12 | 3 | 6 | 1 | 34 |
|  | Percentage of Wins | 23.5% | 11.8% | 35.3% | 8.8% | 17.6% | 2.9% | 100% |
| Wesley | Number of Wins | 6 | 11 | 0 | 2 | 2 | 0 | 21 |
|  | Percentage of Wins | 28.6% | 52.4% | 0% | 9.5% | 9.5% | 0% | 100% |

*Note:* *As percentages have been rounded correctly to one decimal place, totals of percentages may not add to exactly 100%.*

Source: Public Schools Association WA, February 2021, psa.wa.edu.au/records

Use the information in the table to answer the following questions:

(a) Show how the percentage of 35.3% from the Trinity data was calculated. [2]

$$\frac{12}{34} \times 100 = 35.3\%$$

✓ correct numerator

✓ correct denominator
& × 100.

(b) Compare the percentage of total summer sport wins for Aquinas with those for Hale. [3]

Aq: 31.7%                    H: 16.1%

✓ calculates %.

✓ states %. in comparison

✓ makes observation

Aquinas had the most wins with 31.7% with Hale receiving the second most at 16.1%. Hale gained 15.6% less wins than Aquinas.

(c) The total number of wins for each sport over the 40 years may not be exactly 40. Suggest a reason why this number may be:

(i) greater than 40. [1]

Multiple winners - tied.

✓ any valid response

(ii) less than 40. [1]

Sport not existing whole time

✓ any valid response

5. (16 marks)

Finn is a free-range egg farmer. He noted that there appeared to be a relationship between the number of eggs laid and the average maximum daily temperature. The table below gives the number of eggs laid per week ($n$) and the average daily temperature ($t$) in degrees Celsius over a ten-week period starting in July 2020.

| Week | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|------|------|------|------|------|------|------|------|------|------|------|
| $t$ | 19.3 | 19.4 | 19.1 | 19.5 | 19.6 | 19.8 | 19.9 | 20.1 | 20.0 | 19.9 |
| $n$ | 41 | 46 | 40 | 43 | 49 | 53 | 56 | 53 | 57 | 55 |

(a) Add the last three entries to the scatterplot below. [1]

✓ all points plotted correctly.
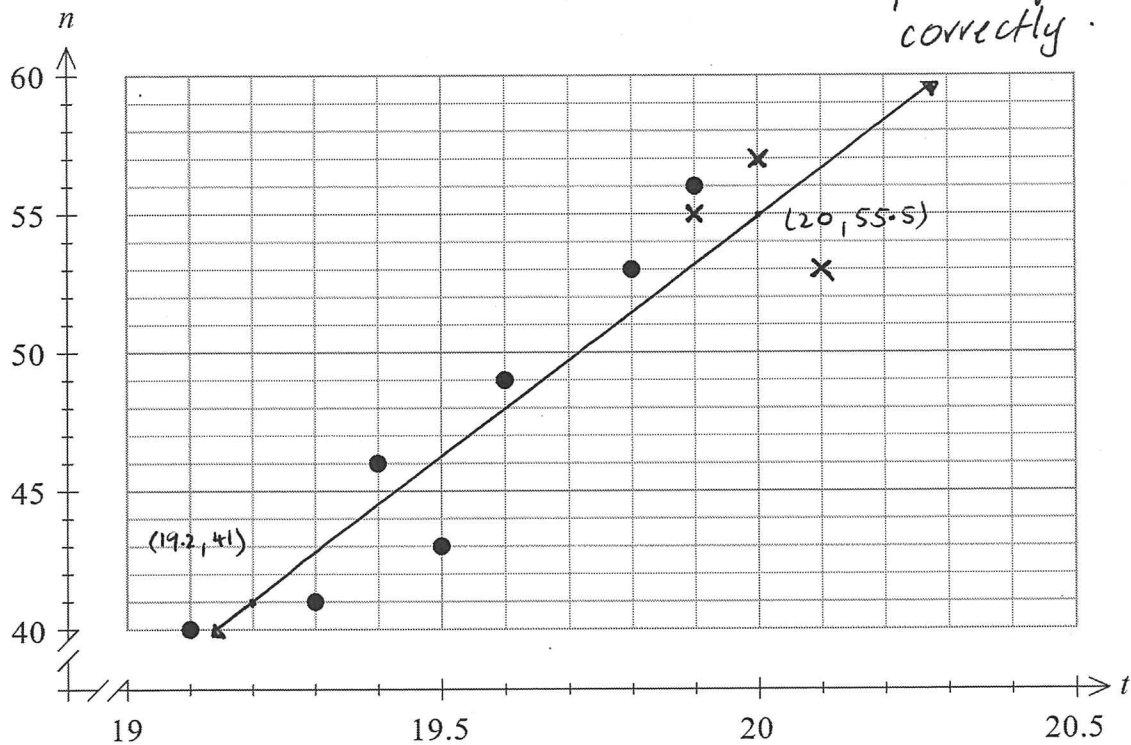


(b) Calculate the correlation coefficient $r_{tn}$ and interpret its value. [2]

$r_{nt} = 0.9338$

strong, positive linear relationship

✓ states r value to 3 d.p. min
✓ strong + positive.

(c) Determine the equation of the least-squares line for the data. [2]

$n = 18.1098t - 306.7378$

✓ correct variables.    ✓ y-int & slope to 2 d.p. min

(d) What would be the expected increase in number of eggs laid for a 0.5 °C increase in temperature? [1]

$0.5 \times 18.1098 = 9.0549$

$\approx 9$ eggs    ✓ correct ans

(e) What percentage of the variation in the number of eggs laid per week is **not** explained by the variation in the temperature in degrees Celsius. [2]

$r^2 = 0.8720$

13 %

100 − 87

*annotations:* ∗ 87% only = no mark.  ✓ calculates $r^2$  ✓ subtracts from 100 to get 13%

(f) Clearly indicating the two points used, add the line of least squares to the scatterplot. [2]

$(19.2, 41)$  $(20, 55.5)$

*annotations:* ✓ indicates points  ✓ draws line correctly for given points.

(g) Finn checks the weather for the following week (week 11) and notes that the average temperature is estimated to be 20.5°C.

(i) How many eggs would Finn predict to be laid in the following week? [1]

$n = 18.1098 (20.5) - 306.7378$

$= 64.51$

$\approx 65$ eggs

*annotations:* ✓ correct ans rounded to nearest integer

(ii) At the end of week 11 Finn notes that he collected 61 eggs during the week. With justification, comment on the reliability of Finn's prediction made for week 11. [2]

Unreliable. Despite the strong correlation coefficient of 0.93 the prediction is extrapolated.

*annotations:* ✓ states unreliable  ✓ justifies w extrapolation.

(h) After estimating the number of eggs for 5 consecutive weeks and comparing these to the actual results, noting all predictions to be within an accuracy of 5 eggs, Finn claims to a neighbouring free-range egg farmer, Francis, that the warmer temperatures cause the chickens to produce more eggs.

(i) Comment on the validity of Finn's statement. [2]

Invalid, correlation does not imply causation.

*annotations:* ✓ Invalid  ✓ explains cause + effect.

(ii) State a non-causal factor that could influence the number of eggs that the chickens produce each week. [1]

coincidence

*annotations:* ✓ any valid response.

6. (6 marks)

Transperth surveyed 300 adult train passengers departing the train at Perth Station on their opinions on train travel. The survey results are tabled below.

| Gender | Positive Opinion | Negative Opinion | Neutral Opinion |
|---|---|---|---|
| Male | 61 | 51 | 23 |
| Female | 61 | 89 | 15 |

(a) A row percentage two-way-table is calculated to analyse the data, explain why? [1]

*The explanatory variable must equate to 100%.*

*✓ valid explanation.*

| Gender | Positive Opinion | Negative Opinion | Neutral Opinion | Total |
|---|---|---|---|---|
| Male | 45% | 38% | 17% | 100% |
| Female | 37% | 54% | 9% | 100% |

(b) Is there an association between gender and opinion on train travel? Justify your answer. [2]

*Yes, "54% of females had a negative opinion on train travel as opposed to 38% of males."*

*✓ Yes    ✓ Valid association w̄ % reference*

~~Opinions of train~~ passengers departing the train at Fremantle Station were also surveyed on their opinions on train travel. This information has been provided on the column graph below.



*≈ 43%*
*≈ 44%*
*≈ 13%*

Legend: ■ Positive Opinion   ▨ Negative Opinion   □ Neutral Opinion
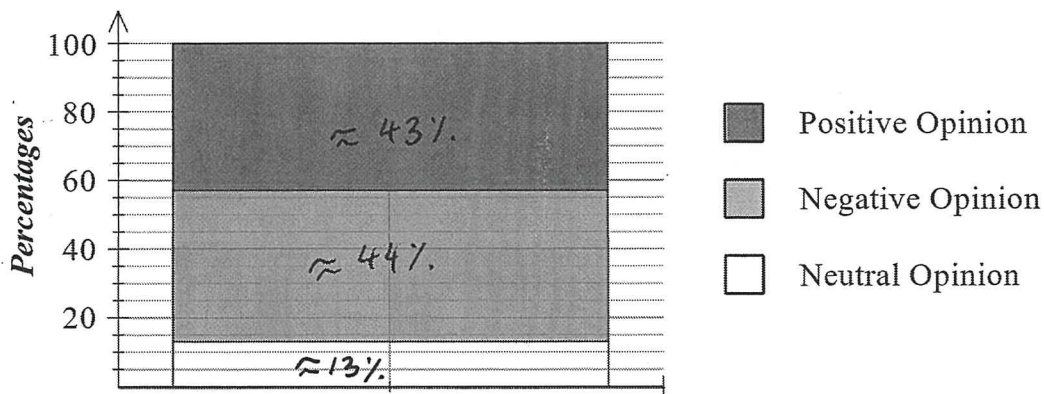
Fremantle

(c) Is there an association between the location of the train station and opinion on train travel. Justify your answer. [3]

*Perth: +ve = 41%    −ve = 47%   Neutral = 13%*

*No association, all opinions have similar percentages regardless of station.*

*✓ calculate perth %'s*

*✓ no association    ✓ justification.*