

## Further Mathematics Practice SAC: Buying a Car

New cars are safer to drive, but they are expensive to buy. This SAC explores the prices of cars of various ages and various makes and attempts to find a satisfactory relationship between the age of a car and its price.

### SECTION 1

Andrew is wondering about whether some vehicles hold their value better than others. He explores the current prices of a number of vehicles that were produced in both 2005 and 2009 and collates the following information including the mean and standard deviation of the prices for each year:

	2009 Car Price (\$)	2005 Car Price (\$)
Ford Falcon	15,500	7,700
Ford Focus	14,400	7,300
Holden Berlina	22,100	9,700
Holden Caprice	38,100	15,400
Holden Commodore	16,830	7,300
Holden Statesman	37,800	12,800
Honda Accord V6	23,500	11,600
Honda Civic	17,300	9,600
Hyundai Elantra	15,400	7,500
Hyundai Sonata	16,600	6,600
Kia Cerato	11,800	7,700
Kia Rio	10,100	6,000
Mazda 3	16,700	11,500
Subaru Liberty	20,300	12,400
Toyota Camry Altise	15,400	7,250
Toyota Corolla	4,900	9,000
Volvo S40	28,000	14,800
Volvo S60	28,100	14,800
Average	20,157.22	9,941.67
Standard deviation	8,099.91	3,080.98

- (1A) Calculate the standardised score for Subaru Liberty for the 2009 data correct to 2 decimal places. (2 marks)

$$z = \frac{x - \mu}{\sigma} = \frac{20300 - 20157.22}{8099.71} = 0.017658 = 0.02$$

- (1B) Comment on the meaning of this standardised score for the Subaru Liberty. (2 marks)

The price of the Subaru is 2% above the average of all cars.

or The price of the Subaru is 0.02 above the average of all cars.

- (1C) Calculate the standardised score for the Subaru Liberty for the 2005 car data correct to 2 decimal places. (2 marks)

$$z = \frac{x - \mu}{\sigma} = \frac{12400 - 9941.67}{3080.98} = 0.8$$

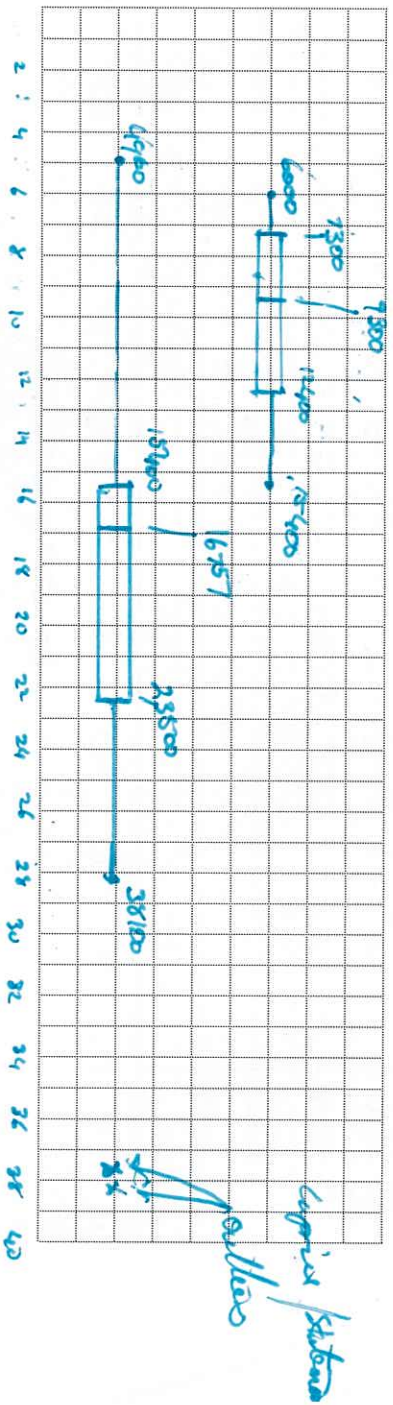
- (1E) Comment on the meaning of any difference in the standardised score for your selected vehicle. (2 marks)

The car price of the Subaru in 2005 was 0.8 above the average car price, whereas in 2009 it was 0.02 above the average car price. This indicates that the average car price for all cars between 2005 and 2009 has increased more in comparison to the price of the Subaru.

- (IG) Calculate the five figure summary for the prices of cars from each of the years 2009 and 2005 and hence complete the table below. (2 marks)

	Min	Q1	Med	Q3	Max
2009 vehicles	4900	15400	16757	23500	38100
2005 vehicles	6000	7300	9300	12400	15400

- (IH) Construct parallel boxplots for the price data from 2005 and 2009 on the grid below. (4 marks)



- (II) Compare the prices for the selected vehicles from each of the years 2005 and 2009. (4 marks)

\* Both positively skewed

\* The median price for 2009 (\$16765) is higher than the median price for 2005 (\$9300)

\* The range of the car price for 2009 (\$33200) is bigger than the range of cars for 2005 (\$9400), and the IQR for 2009 (\$8100) is larger than 2005 (\$5100)

\* The data for 2009 has 2 outliers (more than \$35150). While 2005 has no outliers

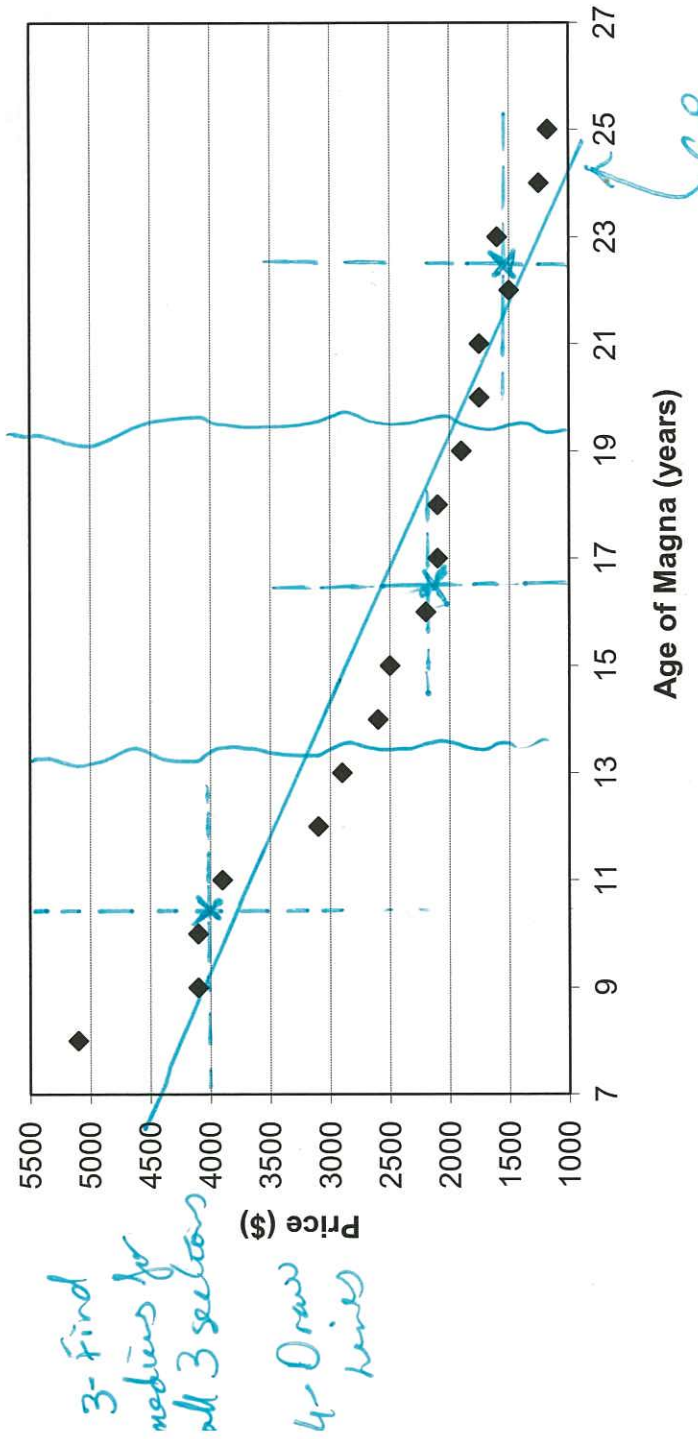
Outliers calc.

2009  $\rightarrow Q_3 + 1.5 \times IQR = 23500 + 1.5 \times 8100 = 35650$   
is the upper boundary, so 37800 and 38100 are outliers.

STEP 1 - Count the Outliers  
 2 - 18 data points 3 groups  
 8/6%

**SECTION 2**

Andrew is interested in buying a Mitsubishi Magna sedan. The scatter plot shows the relationship between the age of a Magna sedan and its price on carsguide.com.au.



(2A) Describe the relationship between the age of the Magna and the price in dollars in terms of strength, direction and form. (3 marks)

Strong Negative linear (1 mark)

(2B) Add the line of best fit by eye to illustrate the trend line of the scatterplot above and comment on the reliability of using this method. (2 marks)

Not accurate as it is a prediction

(2C) The value of the coefficient of determination for the least squares relationship between the price and the age is 0.9035, correct to 4 decimal places. Interpret the meaning of this value in terms of the given data. (2 marks)

90.35% of the variation in the price is explained by the variation in the age of the Magna  
 can be explained by the variation in the age of the Magna  
 Must write the standard form of answers

- (2D) Determine the value of Pearson's product correlation coefficient for this data correct to 3 decimal places. (1)

$$r^2 = 0.9035 \quad \text{so } r = \sqrt{0.9035} \\ = \pm 0.951$$

since this a negative relationship  $r$  is negative so  $r = -0.951$

- (2E) Interpret the meaning of the value calculated in part (2D). (2 marks)

Strong, negative relationship between the Age of the Motor and Price.

- (2F) Given the following information calculate the equation of the least squares regression line for this data and write it below using correct variables. Write the coefficients correct to one decimal place. (4 marks)

mean age of car = 16.5 years  
mean price of car = \$2534.44

standard deviation of car ages = 5.3  
standard deviation of car prices = \$1121.91

$$m = r \times \frac{S_y}{S_x} = -0.951 \times \frac{1121.91}{5.3} = -201.3087 = -201.3$$

$$c = \bar{y} - m\bar{x} = 2534.44 - (-201.3087 \times 16.5) = 5856.0331 = 5856.0$$

rewrite:

$$\text{Price} = -201.3 \times \text{Age of M} + 5856.0$$

- (2G) Explain the meaning in real terms of the gradient in this equation. (2 marks)

The price of the motor decreases its value at the rate of \$201.3 every year on average.

### SECTION 3

Andrew's sister Andrea wants to further explore how the way that the age of a Corolla is related to its price. She looks at the prices of various age Corollas if they are purchased from carsguide.com.au and she gathers the data shown below:

Age of Corolla (years)	Price\$(AUS)
1	15,000
3	14,300
5	10,100
7	7,785
9	6,000
11	3,900
13	3,500
15	3,100
17	2,800
19	2,800
21	2,300
23	1,755
25	1,655
27	1,500
29	1,455
31	1,455

- (3A) Calculate the equation of the least squares regression line for this data and write it below correct to 2 decimal places. (2 marks)

$m = -411.8$ ,  $c = 11551.60$   
equation Price =  $-411.80 \times \text{Age} + 11551.60$

- (3B) Calculate the value of Pearson's product moment correlation coefficient and write it correct to 2 decimal places. (1 mark)

$r = -0.87$

(3C) Interpret the meaning of the value calculated in part (3B).

(2 marks)

Strong negative relationship

(3D) Calculate the value of the coefficient of determination correct to 2 decimal places and interpret the meaning of this value. (3 marks)

$$r^2 = (0.87)^2 = 0.76 \Rightarrow 76\%$$

76% of the variations in the Price of the Corolla can be explained by the variations in the Age of the Corolla.

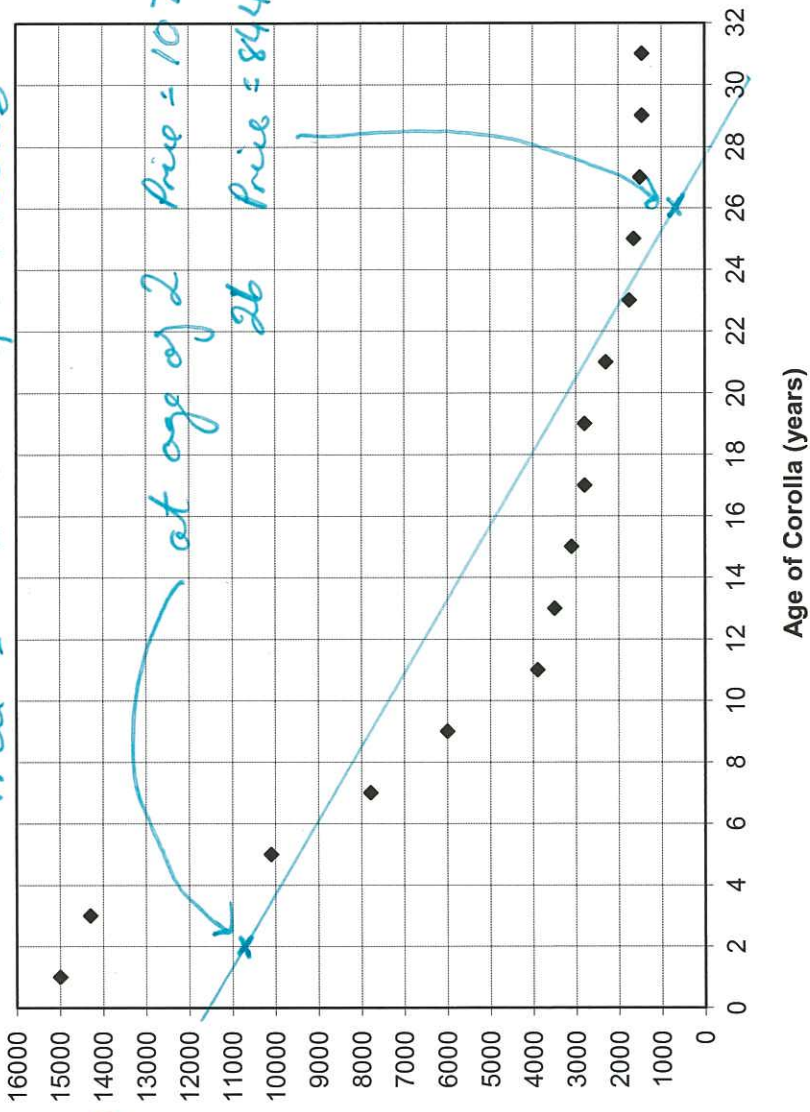
(3E) Use the least squares regression equation to calculate the price of a 10 year old Corolla. (2 marks)

$$\begin{aligned} \text{Price} &: -411.8 \times \text{Age} + 11551.6 \\ &= -411.8 \times 10 + 11551.6 \\ &= -4118 + 11551.6 \\ &= \$7433.60 \end{aligned}$$

(3F) Andrea is looking at a 10 year old Corolla which is advertised at \$6200. Calculate the residual of the price for this vehicle. (2 marks)

$$\begin{aligned} \text{residual} &= \text{actual} - \text{pred.} \\ &= \$6200 - 7433.6 \\ &= -\$1233.60 \end{aligned}$$

(3G) Add the least squares regression line to the scatterplot provided below for the Corolla data. (2 marks)

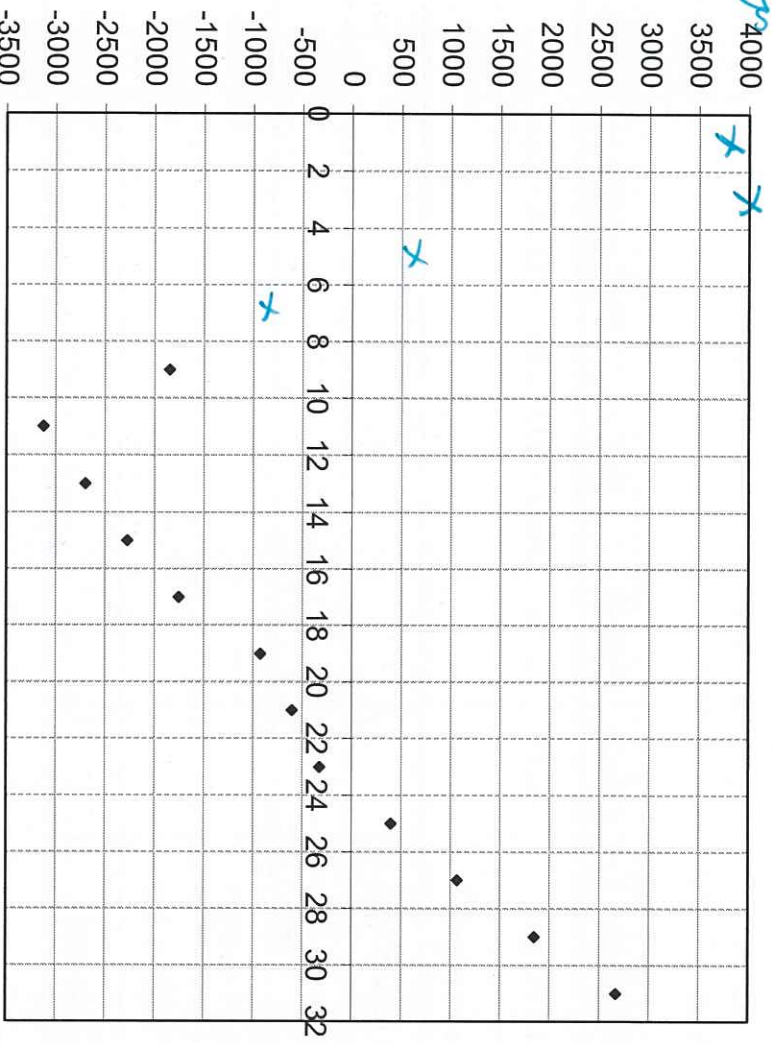


(3H) Complete the table of residuals correct to the nearest dollar for the Corolla data shown below. (2 marks)

Age of Corolla (years)	Residual of price (\$)
1	3860
3	3984
5	607
7	-884
9	-1,845
11	-3,122
13	-2,698
15	-2,275
17	-1,751
19	-927
21	-604
23	-325
25	398
27	1,067
29	1,846
31	2,669



- (3I) The residual plot for the Corolla data is shown below. It is also incomplete. Plot the missing values from the table in part (3H) on this residual plot. (2 marks)



- (3J) What does the residual plot indicate about the relationship between the age of a Corolla and its price? (2 marks)

*The residual plot shows a pattern, this indicates that the relationship of the original data is NOT linear.*

#### SECTION 4

Andrea has come to the conclusion that a transformation may enable her to make a more accurate prediction of the price of the Corolla given its age.

- (4A) Perform each of the transformations listed below and write the values of  $r$ ,  $r^2$  and the coefficients in the equations correct to 2 decimal places. The linear equation has been supplied for your convenience. (18 marks)

Relationship	$r$ value	$r^2$ value	Equation
Price vs Age	-0.87	0.76	Price = 11551.58 - 411.80 × Age
Price vs (Age) <sup>2</sup>	-0.73	0.53	$P = 8533.6 - 10.47 \text{ Age}^2$
Price vs log(Age)	-0.97	0.94	$P = 16536.00 - 10714.01 \times \log(\text{Age})$
Price vs $\frac{1}{\text{Age}}$	0.82	0.68	$P = 2685.58 + 15399.3 \times \frac{1}{\text{Age}}$
(Price) <sup>2</sup> vs Age	-0.77	0.59	Price <sup>2</sup> = 13644823.11 - 5799483.77 × Age
log (Price) vs Age	-0.97	0.94	$\log_{10}(\text{Price}) = 4.12 - 0.04 \times \text{Age}$
$\frac{1}{\text{Price}}$ vs Age	0.99	0.97	Price = -1.44 × 10 <sup>-5</sup> + 2.34 × 10 <sup>-5</sup> × Age

- (4B) State which relationship you consider best and give reasons for your answer. (2 marks)

Price vs Age is the best as it has an  $r^2$  value of 0.97 which is closest to 1.

- (4C) Predict the price of a 10 year old Corolla using the relationship that you consider best correct to the nearest dollar. (2 marks)

$$\text{Price} = -1.44 \times 10^{-5} + 2.34 \times 10^{-5} \times \text{Age}$$

$$= -1.44 \times 10^{-5} + 2.34 \times 10^{-5} \times 10$$

$$= 4533.73 \leftarrow \text{Use solve function}$$

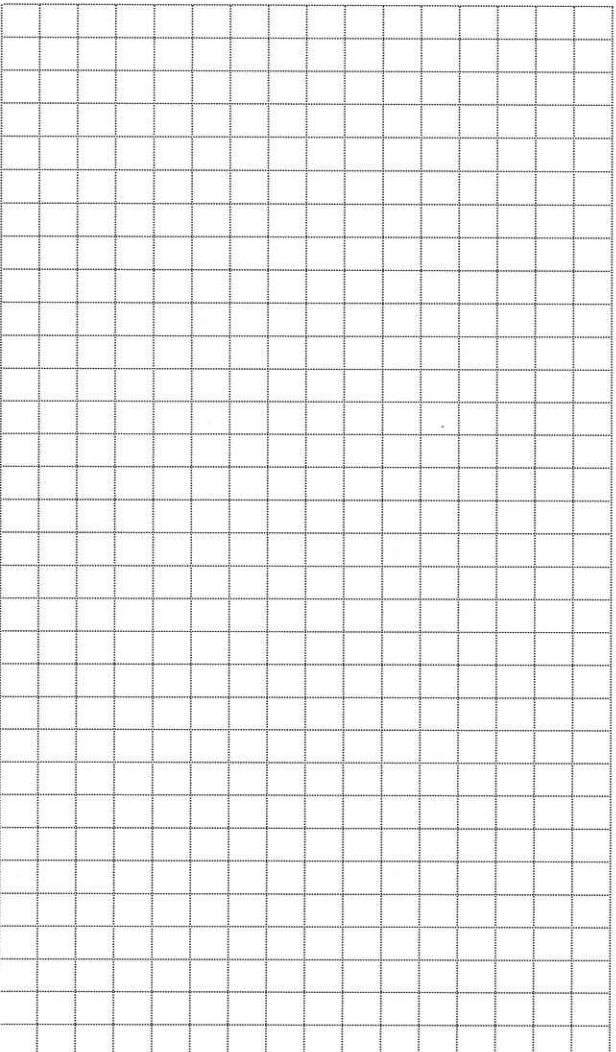
$$\text{Price} = 4554 \quad \text{or } f(10)$$

(4D) Comment on the reliability of your prediction in part (4C).

(2 marks)

The prediction should be reliable as the relationship is strong with an  $r$  of 0.99 and  $r^2$  of 0.97. Moreover, the prediction for the 10<sup>th</sup> year is an interpolator.

(4E) Construct a graph of the relationship that you consider best, showing the scatterplot of the relationship **after** transformation. (3 marks)



(4F) Does this scatterplot after transformation indicate that the transformation was successful? Explain your answer. (2 marks)

Yes, the transformation has been successful because the data has been compressed and concentrated into an almost straight line.

(4G) Calculate the residual for the 10 year old Corolla advertised at \$6200. (2 marks)

$$\begin{aligned} \text{Residual} &= \text{y actual} - \text{y pred} \\ R &= 6200 - 4554 \quad \checkmark \quad \text{from (4C)} \\ R &= 1646 \end{aligned}$$

**SECTION 5**

(5A) Perform a 3-point moving mean smoothing on the following data which represent the sales figures of the Corolla for a 24-month period.

Month	Sales	3-point mean smoothing data
1	34	
2	27	$\frac{34+27+31}{3}$ 30.67
3	31	$\frac{27+31+37}{3}$ 31.67
4	37	$\frac{31+37+41}{3}$ 38.3
5	41	$\frac{37+41+29}{3}$ 35.67
6	29	34
7	32	32.67
8	37	38.67
9	47	40.67
10	38	42
11	41	41
12	44	44
13	47	46.67
14	49	45.67
15	41	47.3
16	52	47
17	48	48
18	44	47
19	49	49.67
20	56	53
21	54	54
23	52	52
24	50	

(5 marks)