Chapter 7 Investigating relationships between two numerical variables: **Assignment**

Name: _____

1 In a study of the length of time it takes students to travel to TAFE in a country town, a researcher collected data from 22 students. The data in the table gives the distance that the students travel (in kilometres) and the time it takes (in minutes).

Distance travelled (km)	*Time taken (mins)*
8	18
8	30
12	15
15	75
20	45
23	60
25	47
40	50
45	80
50	75
50	90
3	5
7	10
8	10
10	10
10	18
15	10
20	30
25	25
30	30
40	42
50	50

a Use your calculator to construct a scatterplot of the *time taken* against the *distance travelled*. In this investigation, *distance travelled* is the explanatory variable.

b Use the scatterplot to describe the association between the variables in terms of strength, direction and form.

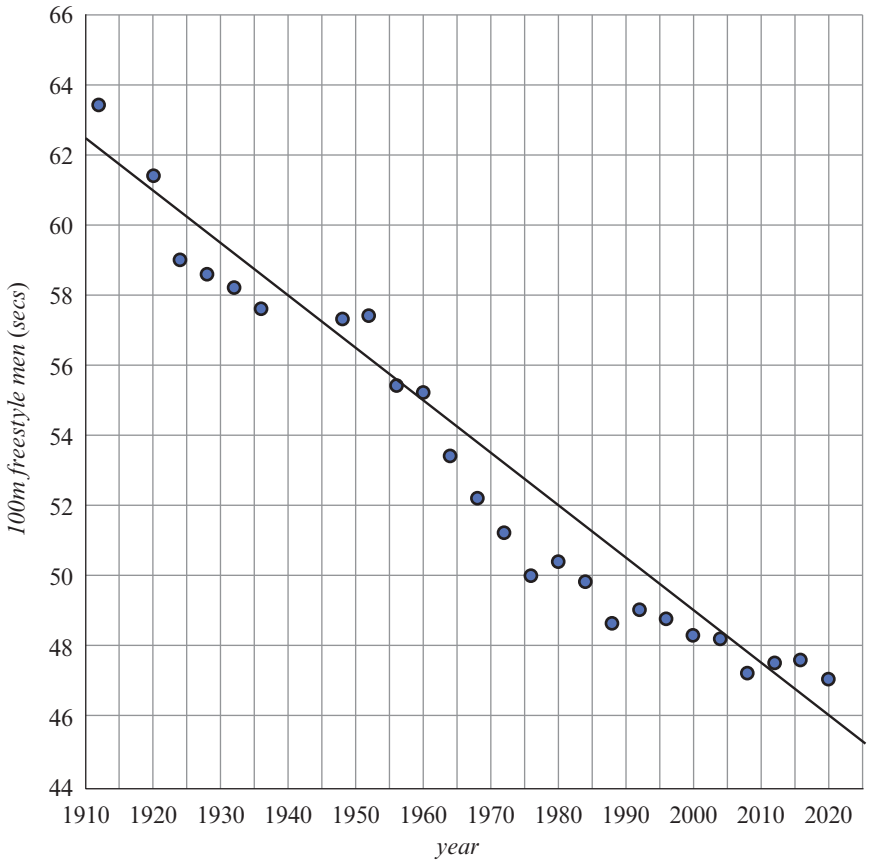
c Determine the correlation coefficient r for this set of data. Write your answer correct to 3 significant figures.

d Use the least squares method to find the equation for the line which will enable *time taken* to be predicted from *distance travelled*. Write the equation in terms of the variables involved and coefficients in the equation correct to 3 significant figures.

e Use the equation of this line to predict the time taken (to the nearest minute) for a student who lives 30 km from the TAFE. In making this prediction, are you interpolating or extrapolating?

2 The Olympic games are generally held every four years. No Olympic games were held between 1936 and 1948 due to World War II. The 2020 games are still referred to as the 2020 games but were held in 2021 due to the Covid-19 pandemic.

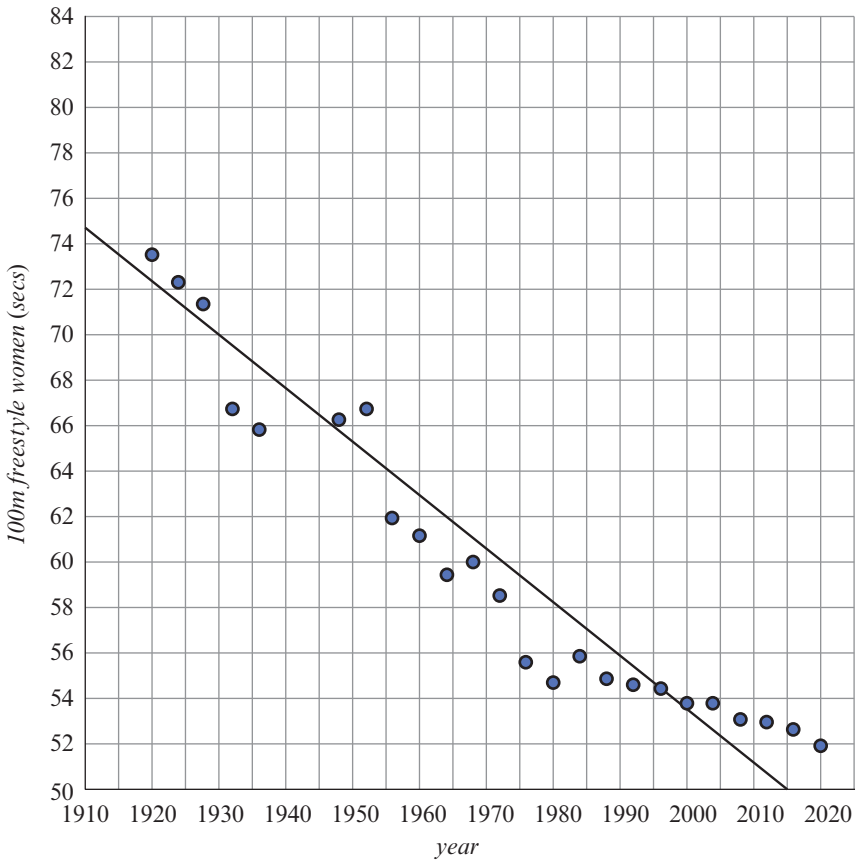
The following scatterplot shows the *time* (in seconds) recorded by the winner of the gold medal in the 100 metre freestyle (men) over the time period 1912–2020.



a **i** Find the equation of the line shown on the scatterplot in terms of *year* and *time*. Round the value of the intercept to the nearest whole number, and the value of the slope to two decimal places.

ii Interpret the value of the slope in terms of the variables *time* and *year*.

b The following scatterplot shows the *time* (in seconds) recorded by the winner of the gold medal in the 100 metre freestyle (women) over the time period 1912–2020.



i Find the equation of the line shown on the scatterplot in terms of *year* and *time*. Round the value of the intercept to the nearest whole number, and the value of the slope to two decimal places.

ii Interpret the value of the slope in terms of the variables *time* and *year*.

c Does your analysis predict that the winning for women will eventually be quicker than the winning time for women? If so, in which Olympic games is this predicted to occur?