

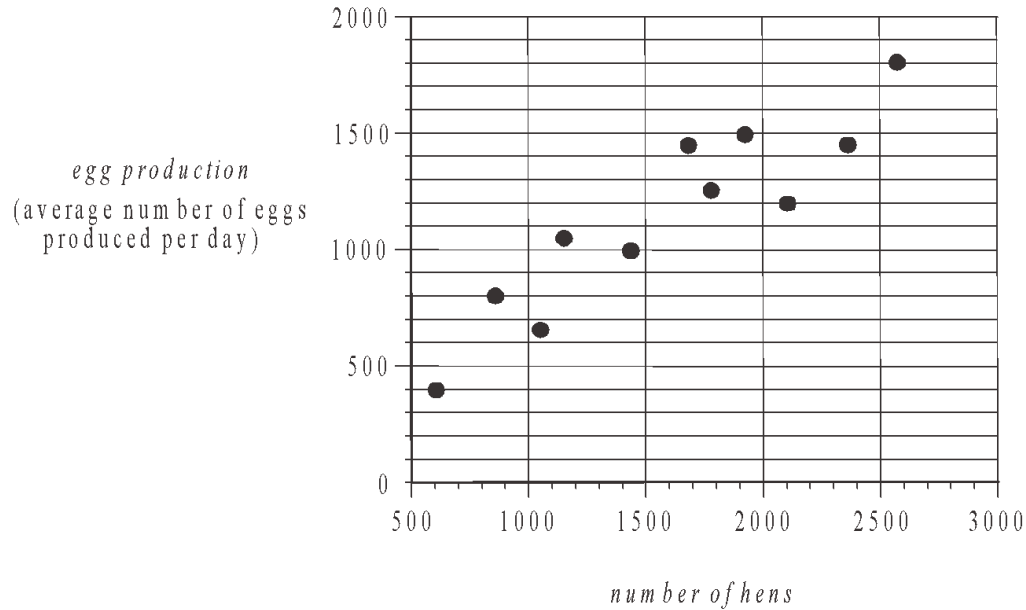
Data Homework – Weeks 3 and 4

Name: _____

Question 1 (8 marks)

The *number of hens* kept at 11 free range egg farms and the *egg production* (average number of eggs produced per day) at each of these farms is recorded.

The scatterplot below displays this data.



The equation of the least squares line is

$$\text{egg production} = 194 + 0.6 \times \text{number of hens}$$

- a.** What is the response variable? 1 mark
-
- b.** Draw the graph of the least squares line on the scatterplot. 1 mark
- c.** Interpret the slope of the least squares line in terms of the variables *egg production* and *number of hens*. 1 mark
-
- d.** At the farm where the *number of hens* is 2 100, there is an *egg production* of 1 200 per day on average. How many eggs does the least squares line overestimate the egg production per day for this farm? 1 mark
-
-

- e. The least squares line equation is used to predict the *egg production* at a farm where the *number of hens* is 4 500.
Give a reason why this prediction might be unreliable. 1 mark

- f. Pearson's correlation coefficient for this data is 0.9182.
Determine the percentage of the variation in *egg production* that can be explained by the variation in the *number of hens* kept at these 11 farms.
Round your answer to two decimal places. 1 mark

The *number of hens* kept at nine **non free-range** egg farms and the *egg production* (average number of eggs produced per day) at each of these farms is recorded.
The data is shown in the table below.

<i>number of hens</i>	1 820	2 450	1 950	2 130	2 210	2 040	1 980	2 380	2 010
<i>egg production</i>	1 440	1 920	1 510	1 630	1 690	1 580	1 560	1 820	1 550

- g. Determine the equation of the least squares line that can be used to predict *egg production* from the *number of hens* kept at a farm.
Write the values of the intercept and slope of the line in the boxes below. Round these values to two decimal places. 2 marks

$$\text{egg production} = \boxed{} + \boxed{} \text{ number of hens}$$