Residuals

Although the r-value indicates the strength and direction of a linear relationship, a lower r-value does not necessarily mean that the linear model should be rejected. Another method of analyzing data is also useful. This involves analyzing the distance the data points are from the line of best fit.

The vertical distance between a data point and the line of best fit is called the **residual value** (or **residual**). In other words, a residual is equal to the **observed value** minus the **predicted value**. It may be calculated for a single point (x_1, y_1) by subtracting the calculated value from the actual value

$$R_1 = y_1 - [a(x_1) + b]$$

where *a* and *b* are the slope and the y- intercept of the line of best fit.

If the model is a good fit, the residuals should be fairly small, and there should be no noticeable pattern. Large residuals or a noticeable pattern are indicators that another model may be more appropriate.

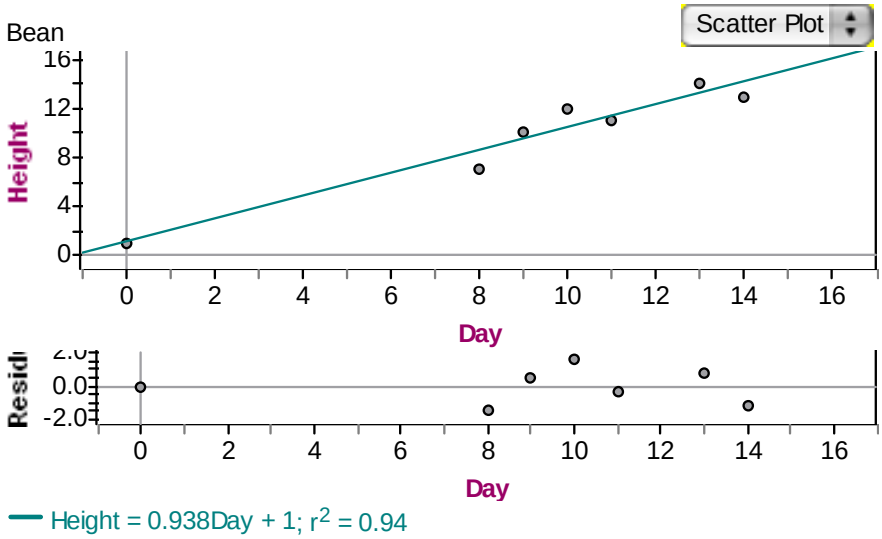
Example 1: One of the science classes is growing a bean plant. The height of the plant is measured every few days. The results are randomly collected.

Bean

	Day	Height
1	0	1
2	10	12
3	8	7
4	13	14
5	9	10
6	11	11
7	14	13

b) Calculate the residuals.

a) Create a scatter plot and find the line of best fit using Fathom.



Refer to the calculations on the board and the completed fathom and excel files.

(0,1): $R_1 = y_1 - (0.938x_1 + 1) = 1 - (0.938 \times 0 + 1) = 1 - (0 + 1) = 1 - 1 = 0$

(10,12): $R_2 = y_2 - (0.938x_2 + 1) = 12 - (0.938 \times 10 + 1) = 12 - (9.38 + 1) = 12 - 10.38 = 1.62$

(8,7): $R_3 = y_3 - (0.938x_3 + 1) = 7 - (0.938 \times 8 + 1) = 7 - (7.504 + 1) = 7 - 8.504 = -1.504$

(13,14): $R_4 = y_4 - (0.938x_4 + 1) = 14 - (0.938 \times 13 + 1) = 14 - (12.194 + 1) = 14 - 13.194 = 0.806$

(9,10): $R_5 = y_5 - (0.938x_5 + 1) = 10 - (0.938 \times 9 + 1) = 10 - (8.442 + 1) = 10 - 9.442 = 0.552$

(11,11): $R_6 = y_6 - (0.938x_6 + 1) = 11 - (0.938 \times 11 + 1) = 11 - (10.318 + 1) = 11 - 11.318 = -0.318$

(14,13): $R_7 = y_7 - (0.938x_7 + 1) = 13 - (0.938 \times 13 + 1) = 13 - (13.132 + 1) = 13 - 14.132 = -1.132$

c) Is the model chosen a good fit based on the residuals?

Since the residuals are small (between -2 and 2) and since they are both positive and negative, then the model chosen (linear model) is a good fit.

Pg. 52: #5-7, 15. (Refer to pg. 421-422 for help with Fathom as needed)