Lesson 2.2.4: Interpreting Residuals

Targets:

1. I can use the line of best fit to predict values for a given data set.
2. I can use residuals to evaluate the accuracy of predictions based on the line of best fit.

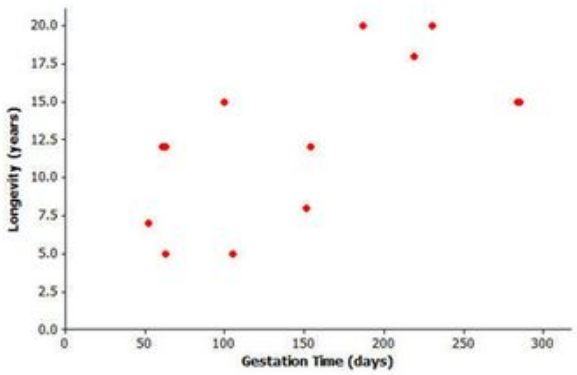
Warm Up

Read through this content and check out the data in the table and the scatter plot. You will be using this information throughout the lesson.

The gestation time for an animal is the typical duration between conception and birth. In other words, gestation time refers to how long the mother is pregnant. The longevity of an animal is the typical lifespan for that animal. In other words, longevity refers to the average number of years an animal lives. The gestation times (in days) and longevities (in years) for 13 types of animals are shown in the table below. The first row says that a Baboon is pregnant for 187 days and that Baboons live for 20 years.

Animal	Gestation Time (days)	Longevity (years)
Baboon	187	20
Black Bear	219	18
Beaver	105	5
Bison	285	15
Cat	63	12
Chimpanzee	230	20
Cow	284	15
Dog	61	12
Fox (Red)	52	7
Goat	151	8
Lion	100	15
Sheep	154	12
Wolf	63	5

Data Source: *Core Math Tools*, www.nctm.org



Practice 1

Finding the equation of the line of best fit relating longevity to gestation time for these types of animal provides the equation to predict longevity. How good is the line? In other words, if you were given the gestation time for another type of animal not included in the original list, how accurate would the line of best fit be at predicting the longevity of that type of animal?

1. Using the Smarter Balanced calculator, find the equation for the line of best fit.
   a. Is this the equation you found: $y = 6.642 + 0.03974x$ where x represents the gestation time (in days) and y represents longevity in years?

2. Graph the line of best fit onto your scatter plot. Your line should be similar to the diagram on my website.

3. Suppose a particular type of animal has a gestation time of 200 days. Approximately what value does the line predict for the longevity of that type of animal?

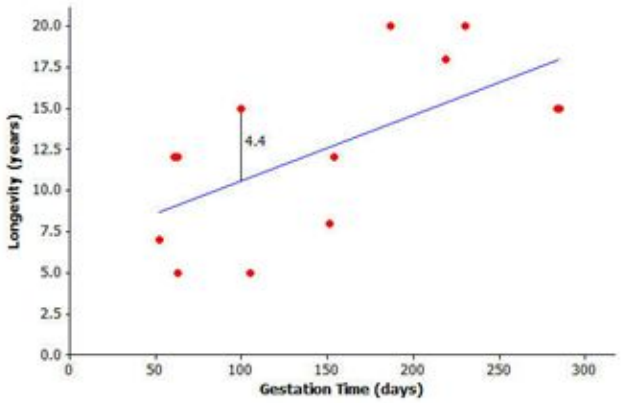
4. Would the value you predicted in question (2) necessarily be the exact value for the longevity of that type of animal? Could the actual longevity of that type of animal be longer than predicted? Could it be shorter?

Practice 2

You can investigate further by looking at the types of animal included in the original data set. Take the lion, for example. Its gestation time is 100 days. You also know that its longevity is 15 years, but what does the line of best fit *predict* for the lion's longevity?
Substituting $x = 100$ days into the equation, you get: $y = 6.642 + 0.03974(100)$or approximately 10.6. The line of best fit predicts the lion's longevity to be approximately 10.6 years.

5. How close is this to being correct? More precisely, how much do you have to add to 10.6 to get the lion's true longevity of 15?

You can show the prediction error of 4.4 years on the graph like this:



Practice 3

6. Let's continue to think about the gestation times and longevities of animals. Let's specifically investigate how accurately the line of best fit predicted the longevity of the black bear. We are comparing the actual longevity to the predicted longevity.

    a. What is the gestation time for the black bear?

    b. Look at the graph. Roughly what does the line of best fit predict for the longevity of the black bear?

    c. Use the gestation time from (a) and the line of best fit $y = 6.642 + 0.03974x$ to predict the black bear's longevity. Round your answer to the nearest tenth.

    d. What is the actual longevity of the black bear?

    e. How much do you have to add to the predicted value to get the actual longevity of the black bear?

    f. Show your answer to part (e) on the graph as a vertical line segment.

Summary of Practice 3

Read the summary of Practice 3 found on the website. Then answer the following questions:

1. What is the formula for calculating residuals?

2. What is a residual?

3. Why is there a difference between the actual value and the predicted value?

Practice 4

Think about what the *actual* longevity of this type of animal might be (the animal with a 270 day gestation).

8. Could it be 30 years? How about 5 years?

9. Judging by the size of the residuals in our table, what kind of values do you think would be reasonable for the longevity of this type of animal? What range of years would you give as a prediction for the longevity of the ocelot?

Continue to think about the gestation times and longevities of animals. The gestation time for the type of animal called the ocelot is known to be 85 days.

- The line of best fit predicts the longevity of the ocelot to be:
    - $y = 6.642 + 0.03974(85) = 10.0$ years

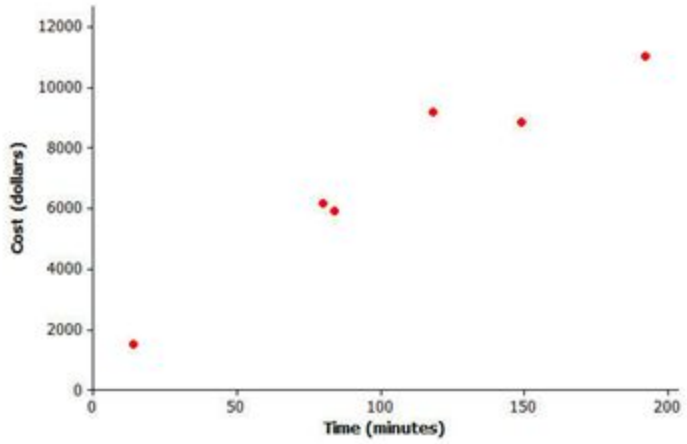
10. Based on the residuals in your table, would you be surprised to find that the longevity of the ocelot was 2 years? Why, or why not? What might be a sensible range of values for the actual longevity of the ocelot?

11. We know that the actual longevity of the ocelot is 9 years. What is the residual for the ocelot?

Exit Ticket

The time spent in surgery and the cost of surgery was recorded for six patients. The results and scatter plot are shown below.

Time (minutes)	Cost ($)
14	1,510
80	6,178
84	5,912
118	9,184
149	8,855
192	11,023



1. Calculate the equation of the line of best fit relating cost to time. (Indicate slope to the nearest tenth and -intercept to the nearest whole number.)

2. Draw the line of best fit on the graph above. (Hint: Substitute $x = 30$ into your equation to find the predicted y- value. Plot the point (30, your answer) on the graph. Then substitute x = 180 into the equation and plot the point. Join the two points with a straightedge.)

3. What does the line of best fit predict for the cost of a surgery that lasts 118 minutes? (Calculate the cost to the nearest cent.)

4. How much do you have to add to your answer to question (3) to get the actual cost of surgery for a surgery lasting 118 minutes? (This is the residual.)

5. Show your answer to question (4) as a vertical line between the point for that person in the scatter plot and the line of best fit.

6. Add a column to your table for the residuals, and then calculate all the residuals for the rest of the data.

7. Suppose that a surgery took 100 minutes.
    a. What does the line of best fit predict for the cost of this surgery?

    b. Would you be surprised if the actual cost of this surgery were $9000? Why or why not?

    c. What range would you give for the cost of a surgery that lasts 100 minutes?