

M112: Modeling with Data: Two-Variable Measurement Data Post-Badge Task

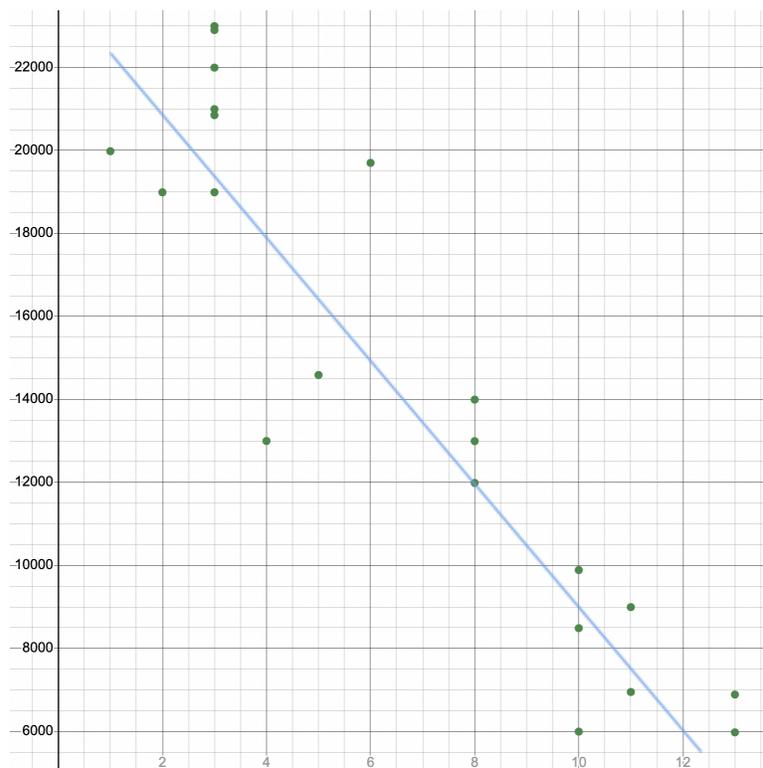
Reflecting on Badge Learning

Now that you have learned more about modeling with two-variable measurement data, respond to the questions that follow. These are the same questions you responded to before the start of this badge.

The Task Illustrative Mathematics, (2016)

Jane wants to sell her Subaru Forester, but doesn't know what the listing price should be. She checks on Facebook Marketplace, finds 22 Subarus listed, and collects data on these cars.

QUESTION 1 Some of the data Jane collected are shown on the graph below. Draw a line that best fits this data. Describe your choice of placement for this line.



I drew this line so that about half of the points are above the line and below the line. This way, the line is as close to as many points as possible.

Teacher Scoring Guide

QUESTION 2 Referring to the graph from Question 1, what could the variables represent about the Subarus that Jane found? Assume that the x-axis represents the independent variable and the y-axis represents the dependent variable.



Answers will vary. Possible student response:

The graphed variables could potentially represent the following:

Independent variable: Age of the vehicle (in years)

Dependent variable: Cost of the vehicle (in dollars)

QUESTION 3 In the graph from Question 1, notice that when the independent variable is 10 there are three dependent variable values. Based on your answer from question 2, what could this mean in context? Does this answer make sense? Why or why not?



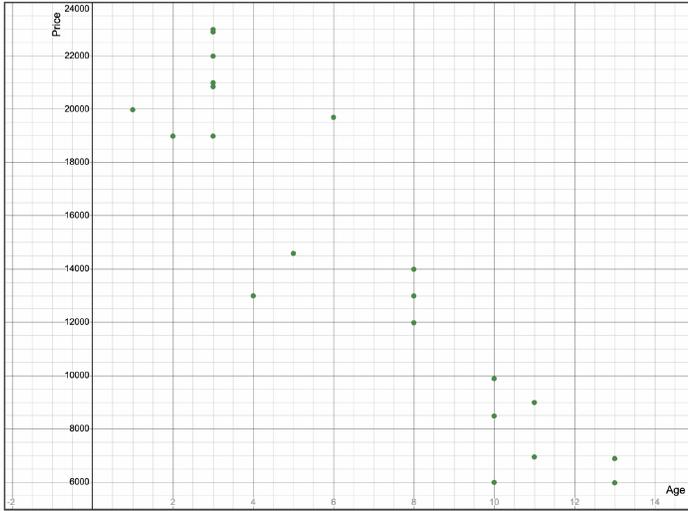
Answers will vary. Possible student response:

For three cars for sale on Facebook Marketplace that are 10 years old, one of them costs \$9800, another costs \$8500, and the last one costs \$6000. This answer makes sense because cars that are 10 years old may not cost the same amount. In this case, the three owners are charging different amounts of money for their 10 year old cars.

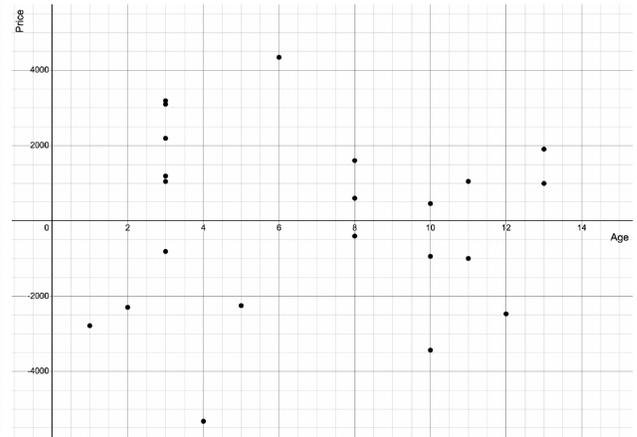
Teacher Scoring Guide

The data Jane collected include age (in years), mileage (in miles), and listed price (in dollars) for the 22 Subarus. Refer to the graphs below for Questions 4-6.

Listed Price (in dollars) vs. Age (in years)



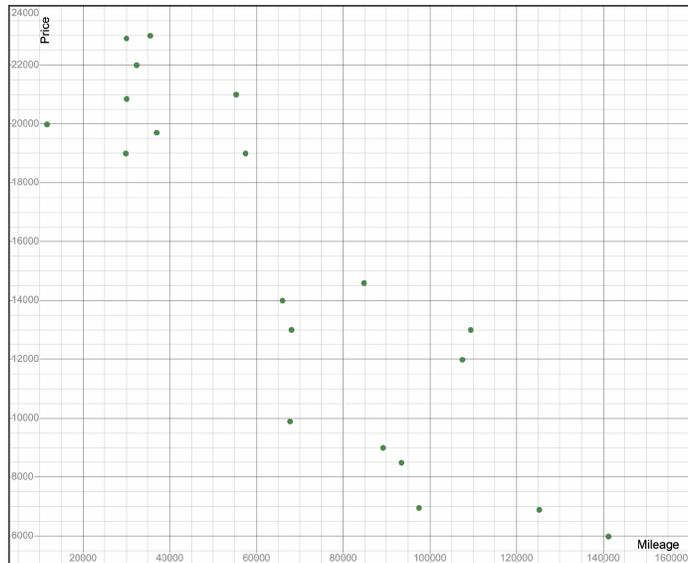
Residual Plot



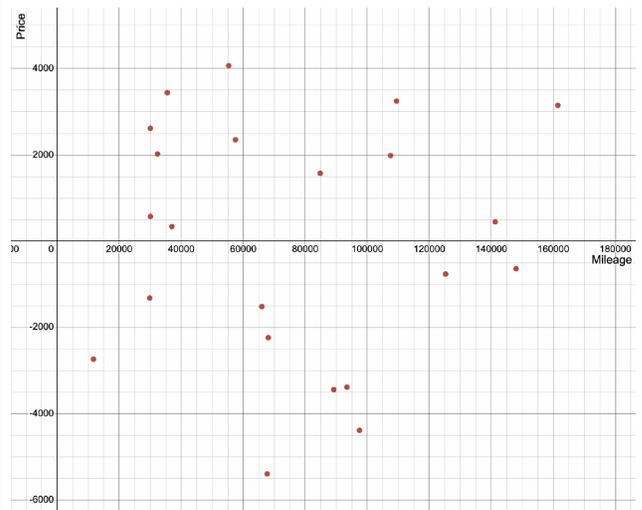
Least Squares Regression Line: $y = -1482.062x + 24247.560$

Correlation Coefficient: $r = -0.9251$

Listed Price (in dollars) vs. Mileage (in miles)



Residual Plot



Least Squares Regression Line: $y = -0.133x + 24265.529$

Correlation Coefficient: $r = -0.8995$

Teacher Scoring Guide

QUESTION 4 Which of the linear models provided (Listed Price vs. Age or Listed Price vs. Mileage) better predicts the listed price? Justify your reasoning.



Both correlation coefficients are fairly close to -1, indicating a strong relationship between the two variables used. Additionally, both residual plots show random scatter/no obvious patterns, which indicates that the linear models created are appropriate.

The regression model for Listed Price vs. Age predicts price slightly better because its correlation coefficient of -0.9251 is slightly closer to -1 than the correlation coefficient for Listed Price vs. Mileage.

QUESTION 5 If Jane's car is 9 years old with 95,000 miles on it, what listing price would you suggest? Explain how you arrived at this price.



$$y = -1482.062(9) + 24247.560 = \$10,909$$

Because we determined that the regression model for Age vs. Listed Price is a slightly better predictor of listed price, we can plug 9 years in for x to produce a prediction. I would suggest a listing price of \$10,909 for Jane's car.

QUESTION 6 Jane's friend Shelbi has a Subaru Forester that is 24 years old with 200,000 miles on it. Would it be appropriate to use either of the models to predict a listing price for Shelbi's car? Why or why not?



The data that Jane collected for Age only shows values from 1 year to 13 years, and 10,000 miles to 150,000 miles for mileage. Because we don't know if the model continues to be linear after 13 years or after 150,000 miles, we cannot reliably predict what Shelbi's car will cost at 24 years old and 200,000 miles.

Teacher Scoring Guide

QUESTION 7 What else did you learn that was not reflected in Questions 1-6?



Answers will vary. Possible student responses:

- Two variables being correlated does not mean that the value of one variable *causes* the other. For example, although we found a strong negative correlation between age and listed price, we cannot say that the older a car is the cheaper it will be. A counterexample to this might be like a vintage classic car that is super expensive.
- The least squares regression line was given to us, but I know how to find this line using technology. And I know that software will find this line by minimizing the sum of the squared residuals.

QUESTION 8 Compare your answers to these questions now to your answers from the earlier attempt. What do you notice about your progress?



Answers will vary. Look for students to explain what they know now compared to what they knew when responding to the pre-test.