Lesson 6: The gradient of a fitted line

Goals

- Describe (orally and in writing) the relationship between two variables using a line fit to data on a scatter plot. Association is used here as correlation has not yet been formally defined.
- Interpret (orally and in writing) points on the scatter plot, including points that do and do not lie on a line fit to the data.
- Interpret (orally and in writing) the gradient of a line fit to data in context.

Learning Targets

- I can use the gradient of a line fit to data in a scatter plot to say how the variables are connected in real-world situations.

Lesson Narrative

In previous lessons, students have been exposed to language like, "As the independent variable increases, the dependent variable tends to decrease." In this lesson they focus on this language in earnest. Students also interpret the gradients of fitted lines in context and identify positive and negative associations of scatter plots without a fitted line shown.

Building On

- Use similar triangles to explain why the gradient m is the same between any two distinct points on a non-vertical line in the coordinate plane; derive the equation $y = mx$ for a line through the origin and the equation $y = mx + c$ for a line intercepting the vertical axis at c.

Addressing

- Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.
- Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line.
- Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the gradient and intercept. For example, in a linear model for a biology experiment, interpret a gradient of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.

Building Towards

- Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the gradient and intercept. For example, in a linear model for a biology experiment, interpret a gradient of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.

Instructional Routines

- Stronger and Clearer Each Time
- Discussion Supports
- Think Pair Share

Student Learning Goals

Let's look at how changing one variable changes another.

6.1 Estimating Gradient

Warm Up: 5 minutes

The purpose of this warm-up is for students to estimate the gradient of a line given points that are close to the line, but not on the line. This prepares students for thinking about the model's fit to data in the rest of the lesson.

Instructional Routines

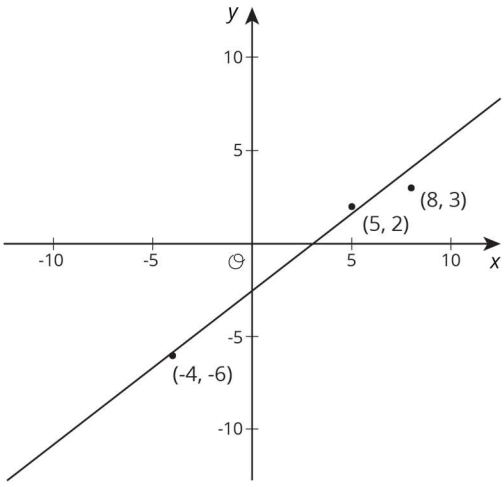
- Think Pair Share

Launch

Arrange students in groups of 2. Give 1 minute of quiet work time followed by 1 minute to discuss their solution with their partner. Follow with a whole-class discussion.

Student Task Statement

Estimate the gradient of the line.



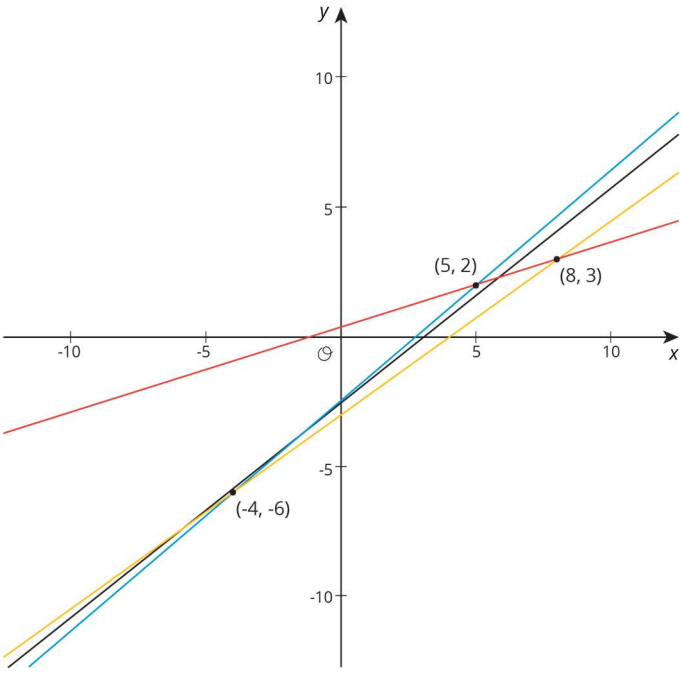
Student Response

The gradient is a little less than $\frac{8}{9}$, which is the gradient of the line segment connecting (-4,-6) with (5,2) since $\frac{2-(-6)}{5-(-4)} = \frac{8}{9}$. The gradient is a little more than $\frac{9}{12} = \frac{3}{4}$, which is the gradient of the line segment connecting (-4,-6) with (8,3) since $\frac{3-(-6)}{8-(-4)} = \frac{9}{12}$. Using the two points in the first quadrant gives a poor estimate of $\frac{1}{3}$.

Activity Synthesis

Poll the class and ask students if their estimated gradient was close to their partner's estimate. Select 2–3 groups who had close estimates to share their solutions and explain their reasoning. Display the graph with the single line and record the students' responses next to the graph for all to see.

If students do not mention that it is better to use points that are far apart rather than close together for estimating the gradient, consider displaying this graph for all to see:



To remind students of previous work, draw a gradient triangle whose horizontal side has a length of 1, demonstrating that the length of the vertical side is equal to the gradient of the line.

6.2 Describing Linear Associations

10 minutes

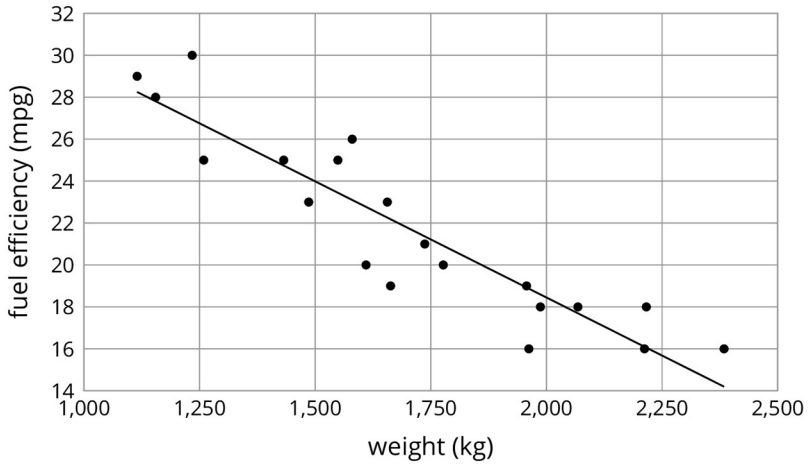
Students practise using precise wording to describe the positive or negative association between two variables given scatter plots of data.

Instructional Routines

- Discussion Supports

Launch

Display the scatterplot for all to see. Remind students that we investigated the relationship between car weight and fuel efficiency earlier.



Ask students how they would describe the relationship between weight and fuel efficiency. After 30 seconds of quiet think time, select 1–2 students to share their responses. (The scatter plot shows that for these cars, as the weight of a car increases, its fuel efficiency decreases.)

Give students 3–5 minutes to construct similar sentences to describe scatter plots of three other data sets followed by a whole-class discussion.

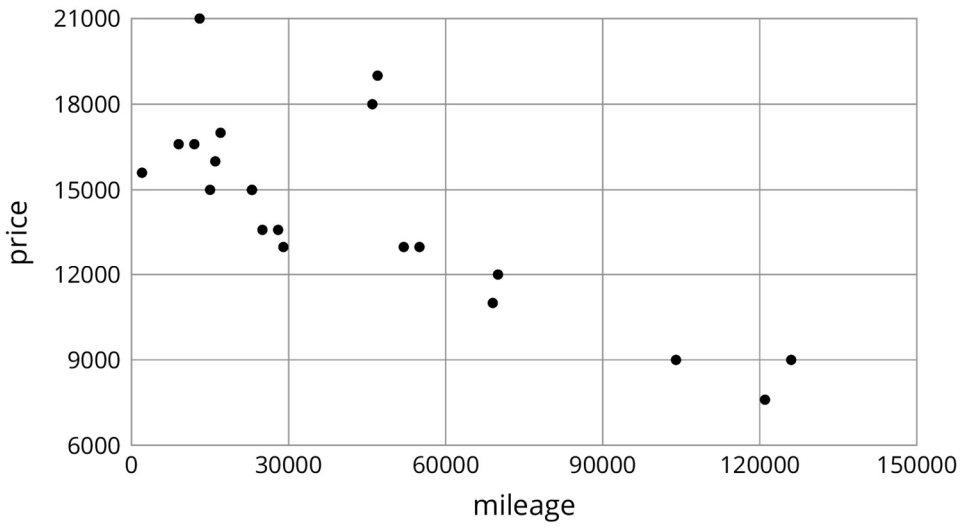
Anticipated Misconceptions

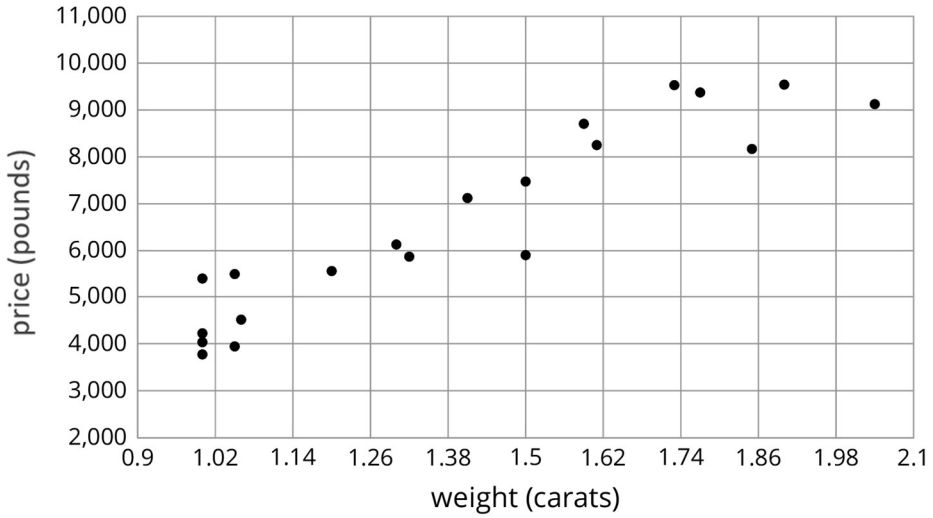
Students may assume they need to use each sentence exactly one time. Let them know it is acceptable to use a sentence more than once and that it is acceptable to not use a sentence.

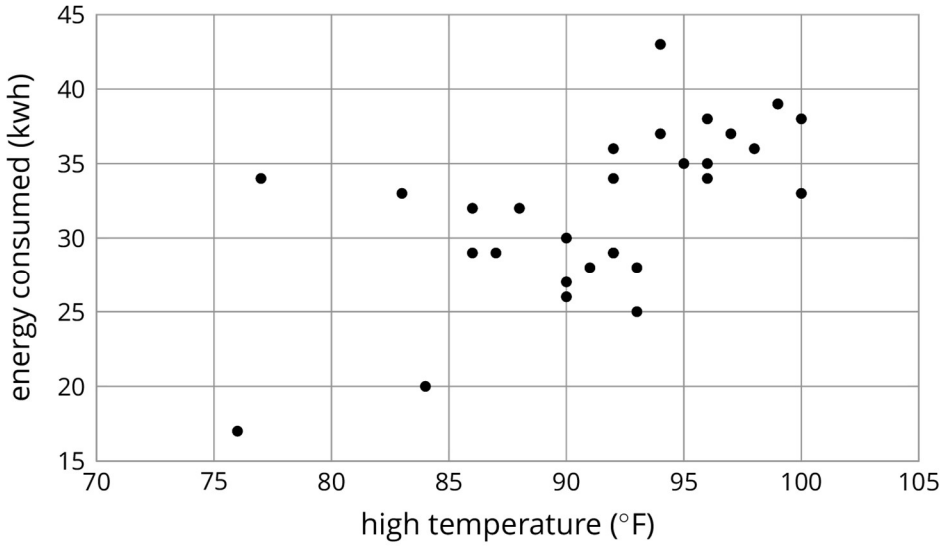
Student Task Statement

For each scatter plot, decide if there is an association between the two variables, and describe the situation using one of these sentences:

- For these data, as _____ increases, _____ tends to increase.
- For these data, as _____ increases, _____ tends to decrease.
- For these data, _____ and _____ do not appear to be related.


21000
18000
15000
12000
9000
6000
price
0
30000
60000
90000
120000
150000
mileage


11,000
10,000
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
price (pounds)
0.9
1.02
1.14
1.26
1.38
1.5
1.62
1.74
1.86
1.98
2.1
weight (carats)


45
40
35
30
25
20
15
energy consumed (kwh)
70
75
80
85
90
95
100
105
high temperature (°F)

Student Response

1. For these data, as car mileage increases, price tends to decrease.
2. For these data, as diamond weight increases, price tends to increase.
3. For these data, as daily high temperature increases, energy consumed tends to increase.

Activity Synthesis

The purpose of the discussion is to talk about trends in data based on the representations in scatter plots.

Consider asking some of the following questions:

- "Is it surprising that the price decreases as mileage increases? Why doesn't mileage predict price perfectly?" (There are other factors, like whether the car has any damage or has any extra features.)
- "Is it surprising that heavier diamonds cost more? Why doesn't weight predict price perfectly?" (There are other factors, like the cut, the clarity, etc.)
- "Is it surprising that energy consumption goes up as the temperature increases?" (No, because of air conditioning.)

For the last scatter plot, highlight the outliers by asking:

- "What might cause more energy consumption on a cool day?" (Laundry, using power tools, etc.)
- "What might cause less energy consumption on a hot day?" (Being gone from home, etc.)

Students may notice that the association between high temperature and energy consumed is more variable than the other situations. There is still a positive association or positive trend, but we would describe the association as "weaker."

Speaking: Discussion Supports. Use this routine to amplify mathematical uses of language to communicate about the relationship between the two quantities represented in each scatter plot. As students describe the trends, press for details by requesting that students challenge an idea, elaborate on an idea, or give an example of their process. Revoice student ideas to model mathematical language use in order to clarify, apply appropriate language, and involve more students.
Design Principle(s): Support sense-making; Optimise output (for explanation)

6.3 Interpreting Gradients

10 minutes

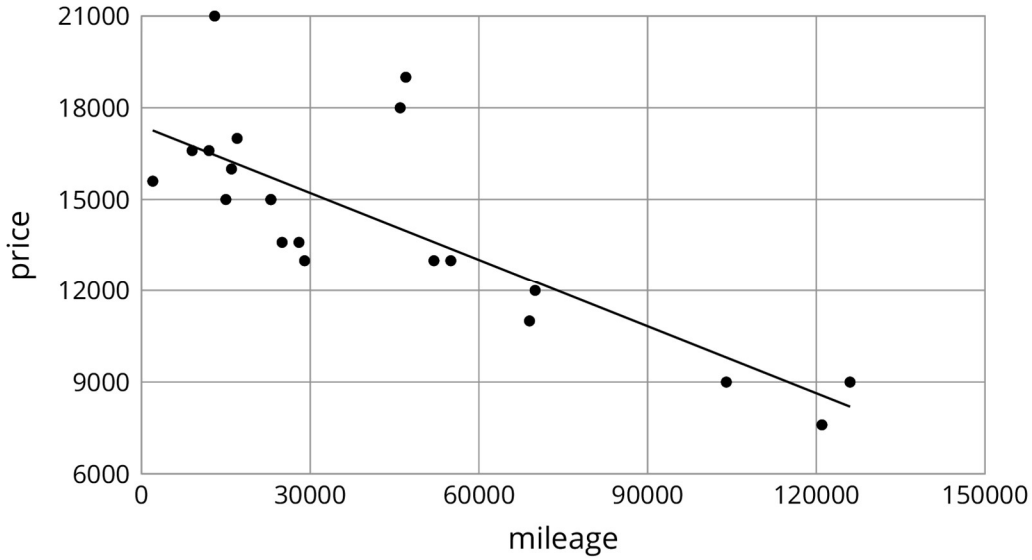
In the previous activity, students noticed trends in the data from the scatter plot. In this activity, the association is made more precise by looking at equations and graphs of linear models for the data to determine the gradient. The numerical value of the gradient is then interpreted in the context of the problem.

Instructional Routines

- Discussion Supports

Launch

Remind students that earlier we looked at the price and mileage of some used cars. We saw that for these used cars, the price tends to decrease as the mileage increases. Display the scatter plot and linear model for the data.



Tell students that an equation for the line is $y = -0.073x + 17\,404.485$. From the equation we can identify the gradient of the line as -0.073. Ask students to think about what that gradient tells us and give quiet think time. Select 1–2 students to share their thinking. (It means that for every increase of one mile, the model predicts that the price of the car will decrease by £0.073.) Tell students that in this activity they will determine what the gradient of the model means for three different sets of data.

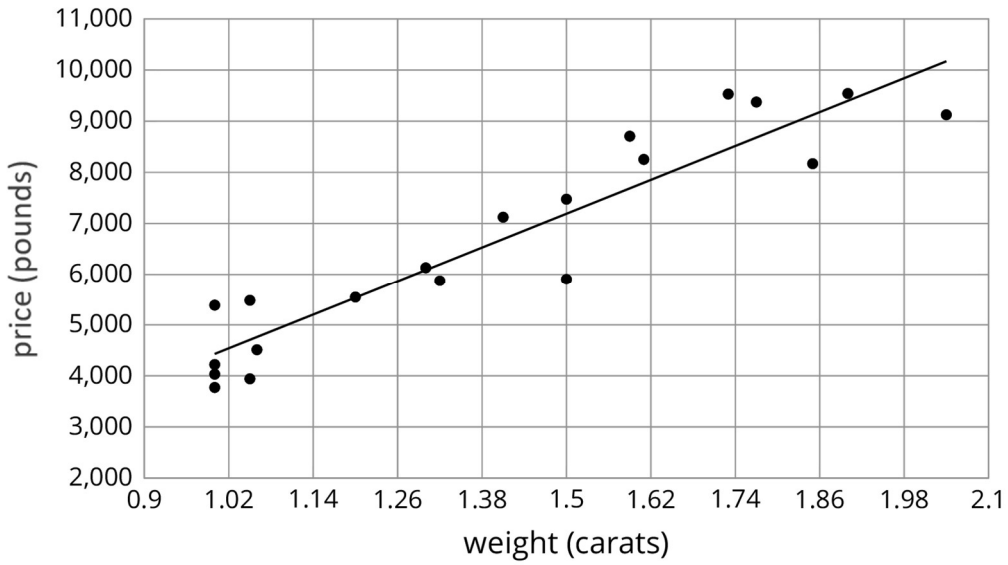
Action and Expression: Internalise Executive Functions. Chunk this task into more manageable parts to support students who benefit from support with organisation and problem solving. For example, present only one scatter plot at a time.
Supports accessibility for: Organisation; Attention Speaking, Writing: Discussion Supports. Provide a sentence frame to scaffold students' language when describing the meaning of gradient in each representation. For example, "If the _____ increases by 1 _____ , the model predicts that _____ (increases/decreases) by _____ ."
Design Principle: Maximise meta-awareness; Optimise output (for explanation)

Student Task Statement

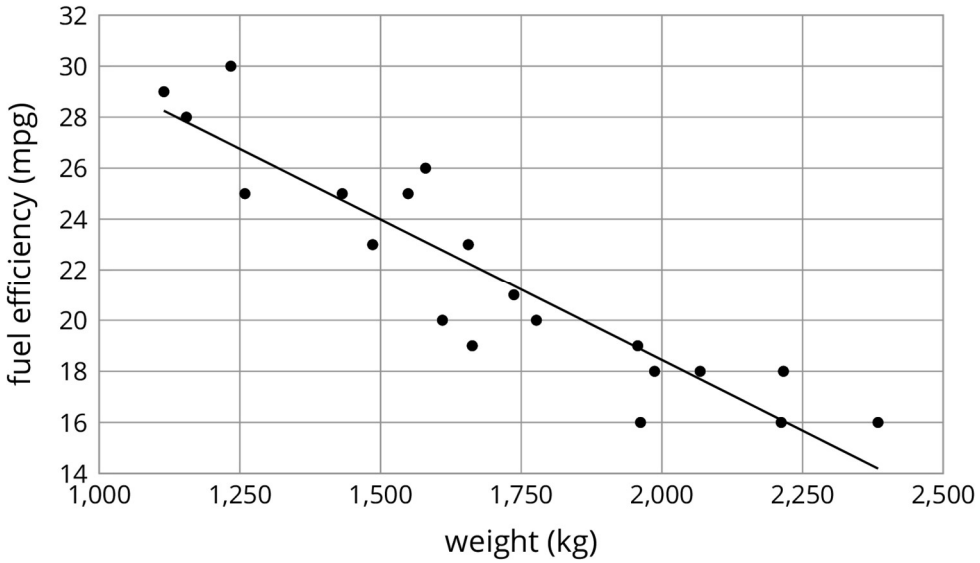
For each of the situations, a linear model for some data is shown.

1. What is the gradient of the line in the scatter plot for each situation?
2. What is the meaning of the gradient in that situation?

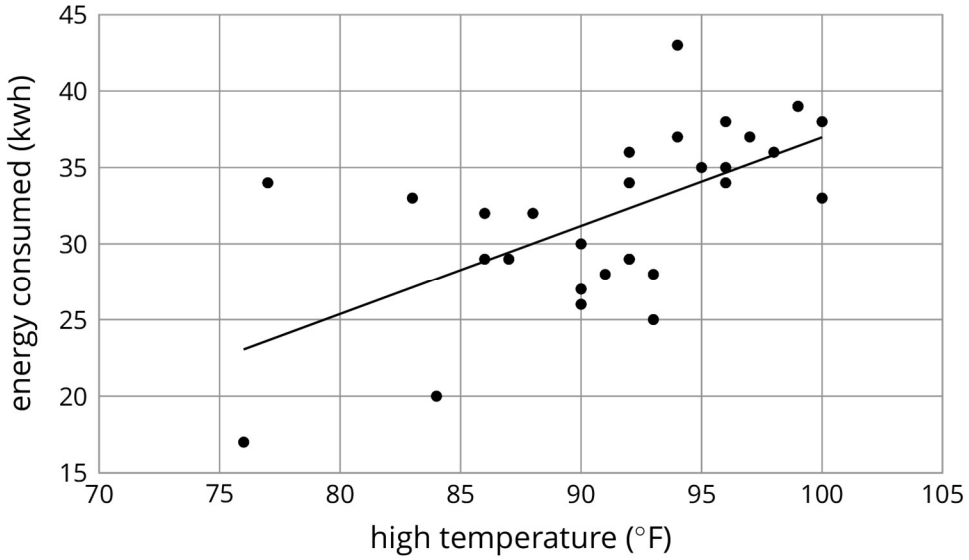
$y = 5\,520.619x - 1\,091.393$



$y = -0.011x + 40.604$



$y = 0.59x - 21.912$



Student Response

For diamond price vs size:

1. The gradient is approximately 5 521.
2. If the weight of a diamond increases by 1 carat, the model predicts that its price increases by £5 521.

For fuel efficiency vs weight:

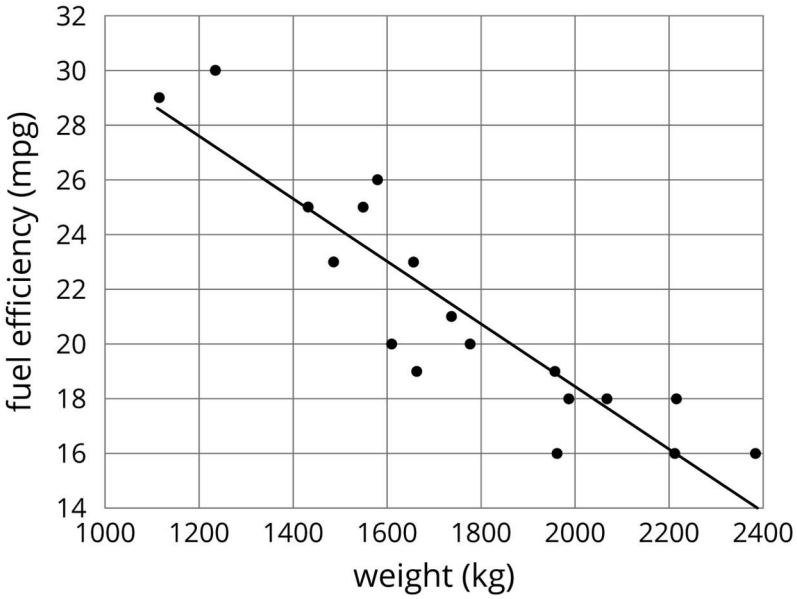
1. The gradient is -0.011.
2. If the weight of a car increases by 1 kg, the model predicts that its fuel efficiency decreases by 0.011 mpg.

Energy consumed vs daily high temperature:

1. The gradient is 0.59.
2. If the daily high temperature increases by 1 degree Fahrenheit, the model predicts that the energy consumed increases by 0.59 kilowatt hour.

Are You Ready for More?

The scatter plot shows the weight and fuel efficiency data used in an earlier lesson along with a linear model represented by the equation $y = \text{-}0.0114x + 41.3021$.



1. What is the value of the gradient and what does it mean in this context?
2. What does the other number in the equation represent on the graph? What does it mean in context?
3. Use the equation to predict the fuel efficiency of a car that weighs 100 kilograms.
4. Use the equation to predict the weight of a car that has a fuel efficiency of 22 mpg.
5. Which of these two predictions probably fits reality better? Explain.

Student Response

1. -0.0114 is the gradient. It means that for every 1 kg increase in the weight of the car, the fuel efficiency decreases by 0.0114 mile per gallon.
2. 41.3021 represents the value of the y-coordinate when the x-coordinate is zero. In other words, the y-intercept of the graph. It means that a car that weighs 0 kg would have a fuel efficiency of about 41 miles per gallon.
3. 40.16 mpg since $-0.0114 \times 100 + 41.3021 \approx 40.16$.
4. 1 693 kg since $(22 - 41.3021) \div -0.0114 \approx 1\,693$.
5. Answers vary. Sample response: The weight of the car with 22 mpg probably fits reality better. A 100 kilogram car is so small that it probably doesn't fit with the trend of the data we are using to create the linear model.

Activity Synthesis

The purpose of this discussion is to develop a quantitative sense of trends based on linear models of the data.

Consider asking some of the following questions:

- "What was the easiest way to find the gradient for each situation?" (The coefficient of the x-coordinate in the equation.)
- "Do the answers for the meaning of the gradients make sense in the contexts of the problem?" (Yes. Diamonds are usually expensive, so a 1 carat increase is likely to cost a lot more. A 1 kilogram difference in the weight of a car is not very much, so it may lower the gas mileage, but not by much. A 1 degree temperature increase is not significant, so it should not need a lot more energy to cool off a building, but some would be needed.)
- "What is the difference between a positive gradient and negative gradient in your interpretations?" (A positive gradient means both variables increase together. A negative gradient means that when one variable increases, the other decreases.)
- "The model for energy consumption and temperature predicted a 0.59 kilowatt hour increase in energy consumption for every 1 degree increase in temperature. Estimate how much the temperature would need to increase to raise energy consumption by 6 kilowatt hours." (A little more than 10 degrees.)

6.4 Positive or Negative?

10 minutes

This activity returns to scatter plots without linear models given. Students determine whether the data seems to have a linear association or not. If it does, students are asked to decide whether the variables have a positive or negative association.

Instructional Routines

- Stronger and Clearer Each Time

Launch

Tell students that while some data sets have a *linear association*, others do not. A linear association is present when the points in a scatter plot suggest that a linear model would fit the data well. Some data sets have a *non-linear association* when the scatter plot suggests that a non-linear curve would fit the data better. Still other data sets have *no association* when the data appears random and no curve would represent the data well.

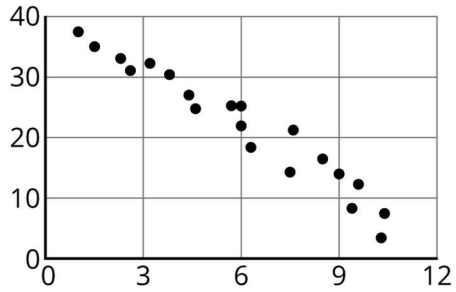
In this activity, students first need to identify if data has a linear association or not and, if it does, what type of gradient a linear model of the data would have.

Representation: Develop Language and Symbols. Create a display of important terms and vocabulary. Invite students to suggest language or diagrams to include that will support their understanding of linear association, no association, positive association, and negative association.
Supports accessibility for: Memory; Language

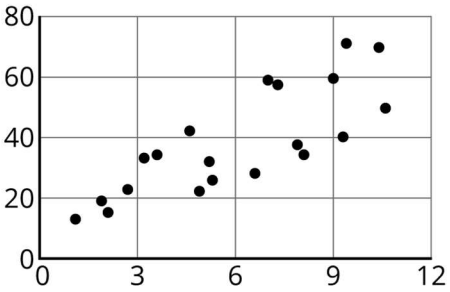
Student Task Statement

1. For each of the scatter plots, decide whether it makes sense to fit a linear model to the data. If it does, would the graph of the model have a positive gradient, a negative gradient, or a gradient of zero?

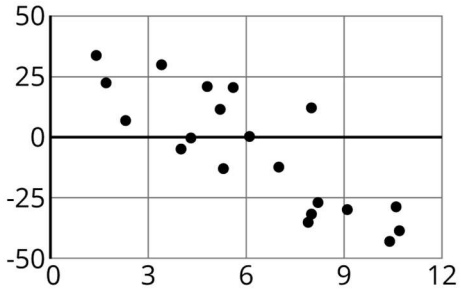
A



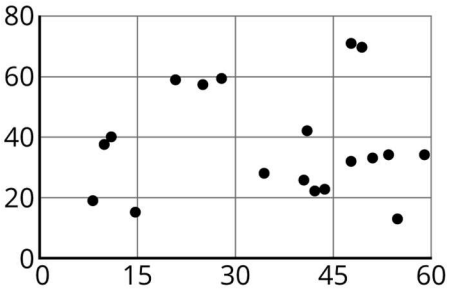
B



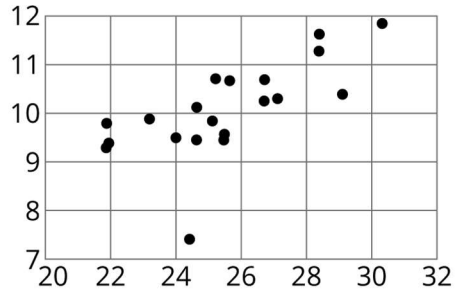
C



D



E



2. Which of these scatter plots show evidence of a positive association between the variables? Of a negative association? Which do not appear to show an association?

Student Response

1. A: A linear model would fit and have a negative gradient.

   B: A linear model would fit and have a positive gradient.

   C: A linear model would fit and have a negative gradient.

   D: A linear model would not fit very well.

E: A linear model would fit and have a positive gradient, even though it has an outlier.

2. B and E have a positive association. A and C have a negative association. D has no apparent association.

Activity Synthesis

The purpose of this discussion is to solidify understanding of trends in scatter plots and look for associations in the data.

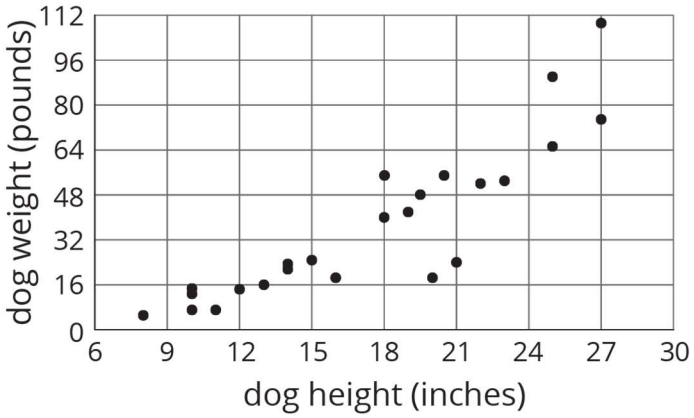
Some questions for discussion:

- "What strategy did you use to determine if the data had a linear association?" (If a single line was fairly close to all of the points, I said it has a linear association.)
- "Would you be able to notice a linear association from the data in a table?" (It might be possible by finding the gradient between each pair of points and seeing if they are close, but would be very difficult. Using a scatter plot is much easier.)
- "What does it mean to have a positive association?" (When one of the variables increases, the other does, too.)
- "How did you determine if the data had a positive association?" (When the points generally go from the bottom left to the top right, there should be a positive association.)
- "What are other pairs of things that you conjecture would have a negative, linear association?"

Writing: Stronger and Clearer Each Time. Use this routine to give students a structured opportunity to revise their response to the last question. Ask each student to meet with 2–3 other partners in a row for feedback. Provide students with prompts for feedback that will help students strengthen their ideas and clarify their language (e.g., "Why do you think it's a positive association?", "How did you determine the association between the two quantities?", and "How could you incorporate an example from the data set to support your ideas?", etc.). Students can borrow ideas and language from each partner to strengthen their final explanation.
Design Principle(s): Optimise output (for explanation)

Lesson Synthesis

Display the scatter plot for all to see.



To highlight the main ideas from the lesson about the meaning of the gradient of a fitted line, ask:

- "How would you describe the trend in the scatter plot?" (As a the height of a dog increases, its weight tends to increase.)
- "When there is an association between two variables, how can we tell if it is a positive association or a negative one?" (If the dependent variable tends to increase as the independent variable increases, it is a positive association, and if the dependent variable tends to decrease, it is a negative association. Also, a line that is a good fit for the data will have a positive gradient (negative gradient, respectively).)
- "The gradient of a line that models the data is 4.25. What does that tell us about the dogs?" (For every 1 inch increase in dog height, the predicted weight increase is 4.25 pounds.)

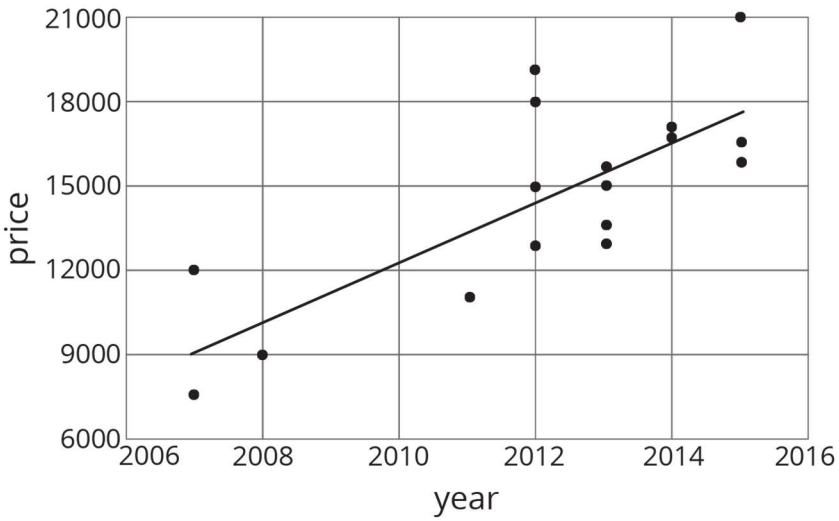
6.5 Trends in the Price of Used Cars

Cool Down: 5 minutes

Students use a scatter plot of data representing used car prices and their model year to determine the sign of the gradient of a linear model. Then students interpret this gradient in the context of the data.

Student Task Statement

Here is a scatter plot that shows the years when some used cars were made and their prices in 2016, together with the graph of a linear model for the relationship between year and price.



1. Is the gradient positive or negative?

2. Which of these values is closest to the gradient of the linear model shown in the scatter plot?

   - 1 000
   - 3 000
   - -1 000
   - -3 000

3. Use the value you selected to describe the meaning of the gradient in this context.

Student Response

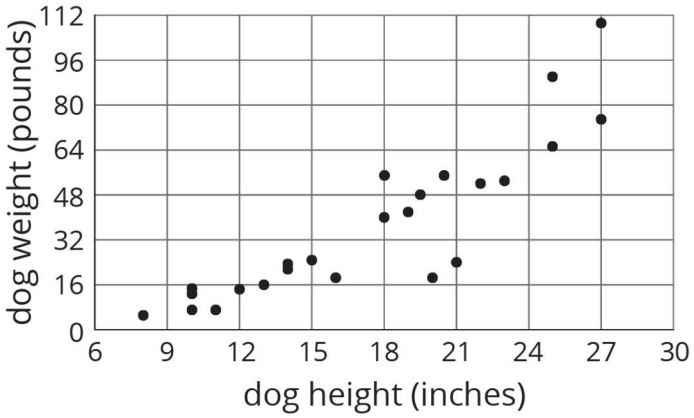
1. The gradient is positive, because as the year of the car increases, the price tends to increase.

2. 1 000

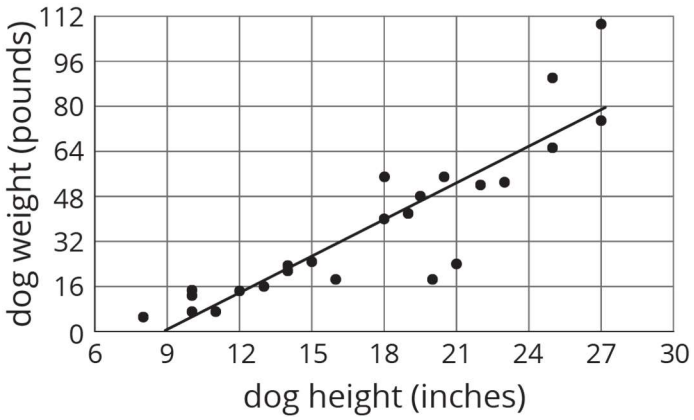
3. The model predicts that when a car is made 1 year later, the price is £1 000 higher.

Student Lesson Summary

Here is a scatter plot that we have seen before. As noted earlier, we can see from the scatter plot that taller dogs tend to weigh more than shorter dogs. Another way to say it is that weight tends to increase as height increases. When we have a positive association between two variables, an increase in one means there tends to be an increase in the other.

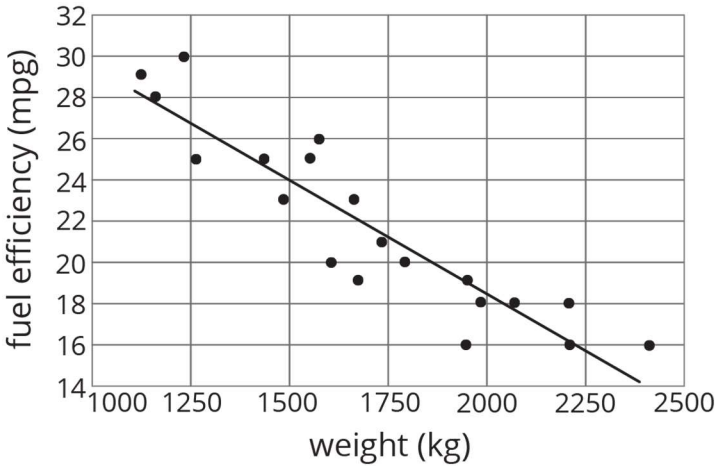


We can quantify this tendency by fitting a line to the data and finding its gradient. For example, the equation of the fitted line is $w = 4.27h - 37$ where h is the height of the dog and w is the predicted weight of the dog.



The gradient is 4.27, which tells us that for every 1 inch increase in dog height, the weight is predicted to increase by 4.27 pounds.

In our example of the fuel efficiency and weight of a car, the gradient of the fitted line shown is -0.01.

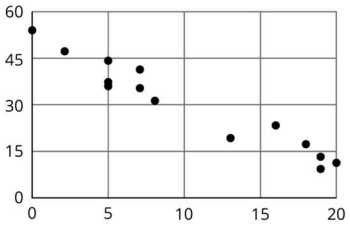


This tells us that for every 1 kilogram increase in the weight of the car, the fuel efficiency is predicted to decrease by 0.01 miles per gallon. When we have a negative association between two variables, an increase in one means there tends to be a decrease in the other.

Lesson 6 Practice Problems

Problem 1 Statement

Which of these statements is true about the data in the scatter plot?



a. As x increases, y tends to increase.

b. As x increases, y tends to decrease.

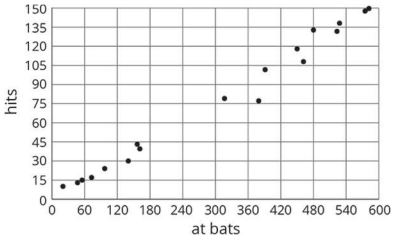
c. As x increases, y tends to stay unchanged.

d. x and y are unrelated.

Solution B

Problem 2 Statement

Here is a scatter plot that compares hits to at bats for players on a baseball team.



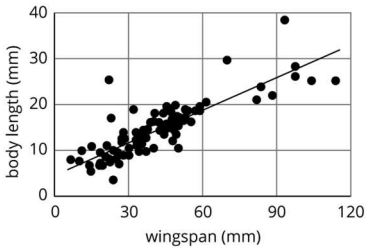
Describe the relationship between the number of at bats and the number of hits using the data in the scatter plot.

Solution

As the number of at bats increases, the number of hits also increases.

Problem 3 Statement

The linear model for some butterfly data is given by the equation $y = 0.238x + 4.642$. Which of the following best describes the gradient of the model?



a. For every 1 mm the wingspan increases, the length of the butterfly increases 0.238 mm.

b. For every 1 mm the wingspan increases, the length of the butterfly increases 4.642 mm.

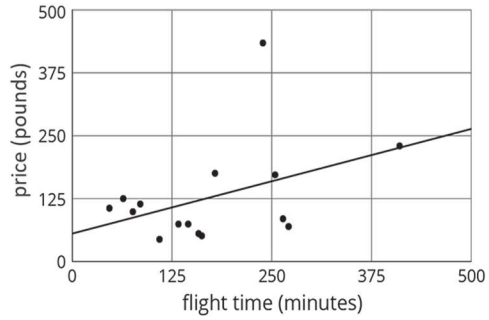
c. For every 1 mm the length of the butterfly increases, the wingspan increases 0.238 mm.

d. For every 1 mm the length of the butterfly increases, the wingspan increases 4.642 mm.

Solution A

Problem 4 Statement

Nonstop, one-way flight times from O'Hare Airport in Chicago and prices of a one-way ticket are shown in the scatter plot.



a. Circle any data that appear to be outliers.

b. Use the graph to estimate the difference between any outliers and their predicted values.

Solution

a. The point at (239, 436) appears to be an outlier.

b. This point represents a destination that costs around £250 more than the model predicts for its flight time.

Problem 5 Statement

Solve: $\begin{cases} y = -3x + 13 \\ y = -2x + 1 \end{cases}$

Solution

(12,-23)