Unit 5, Lesson 21

Describing Trends in Scatter Plots

Goals

- Draw a linear model to fit data in a scatter plot, and describe (in writing) features of a line that fit data well.
- Draw a linear model to fit data in a scatter plot, and describe (in writing) features of a line that fits data well.

Learning Targets

- I can draw a line to fit data in a scatter plot.
- I can say whether data in a scatter plot has a positive or negative association (or neither).

Lesson Narrative

In this lesson, students are introduced to the terms **positive association** and **negative association**. They use fitted lines to help them understand this language and tie it back to their work in an earlier unit on linear relationships. They start to use language to describe trends like, "Cars made in a later year tend to have a higher price." They evaluate the goodness of fit of lines for a given scatter plot (note that this is done informally and in later courses, students will revisit more quantitative measures of fitness) and begin to draw their own lines to fit data in a scatter plot (MP5).

Teacher Notes for IM 6–8 Math Accelerated v.360
The *Activity Narrative* mentions work with linear relationships in an earlier unit. That work now occurs earlier in the current unit.

Standards

Addressing **8.SP.A.1, 8.SP.A.2**

Instructional Routines

- 5 Practices
- MLR1: Stronger and Clearer Each Time
- MLR2: Collect and Display
- Notice and Wonder
- Which Three Go Together?

Required Materials

Materials to Gather

- Dried linguine pasta: Activity 2
- Straightedges: Activity 2

Required Preparation

Activity 2:

For the digital version of the activity, acquire devices that can run the applet.

Lesson:

Each student needs one strand of pasta. Have extra available in case the strands break.

Student Facing Learning Goals

Let's look for associations between variables.

21.1 Which Three Go Together: Scatter Plots

5 min

Warm-up

Activity Narrative

This warm-up prompts students to compare four scatter plots. It gives students a reason to use language precisely (MP6). It gives the teacher an opportunity to hear how students use terminology and talk about characteristics of the items in comparison to one another.

It also introduces students to positive and negative associations by comparing scatter plots with best-fit lines.

Standards

Addressing **8.SP.A.1, 8.SP.A.2**

Instructional Routines

- Which Three Go Together?

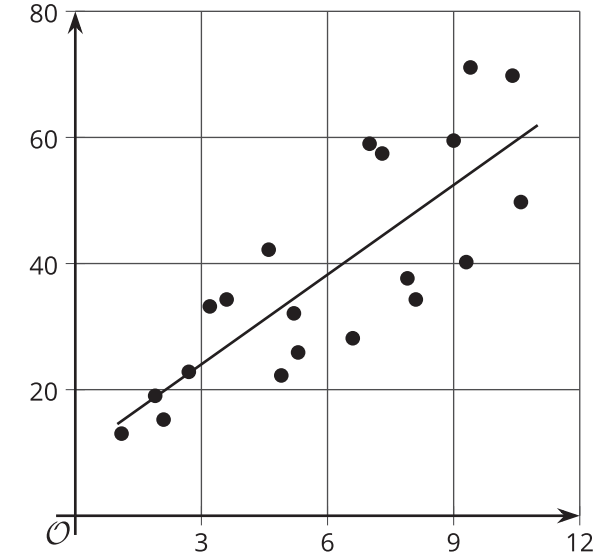
Launch

Arrange students in groups of 2–4. Display the scatter plots for all to see. Give students 1 minute of quiet think time and ask them to indicate when they have noticed 3 plots that go together and can explain why. Next, tell students to share their response with their group, and then together find as many sets of three as they can.

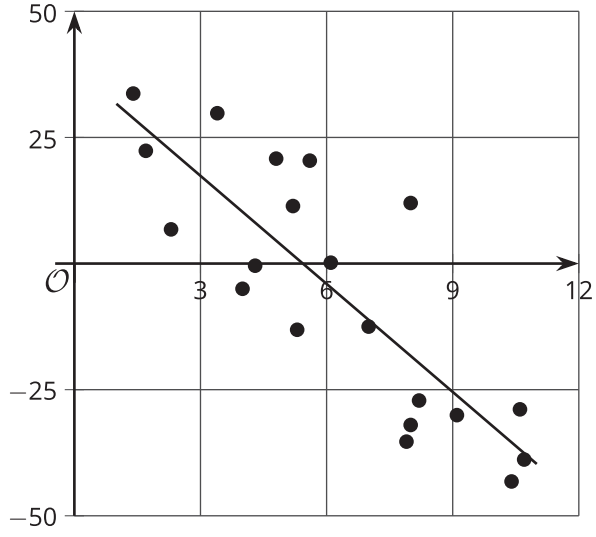
Student Task Statement

Which three go together? Why do they go together?

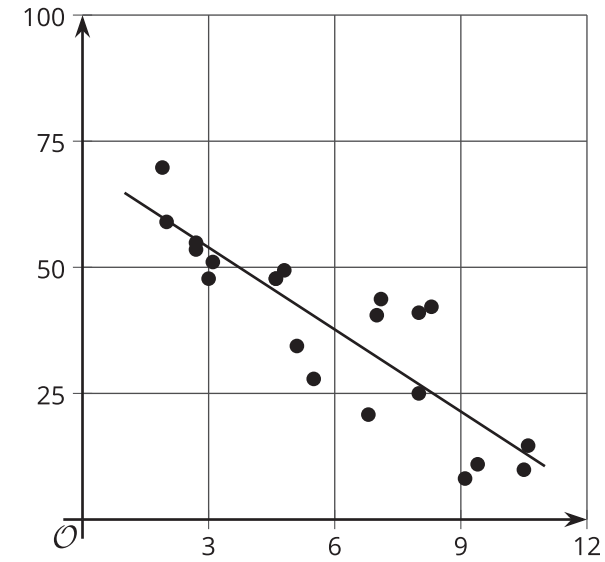
A



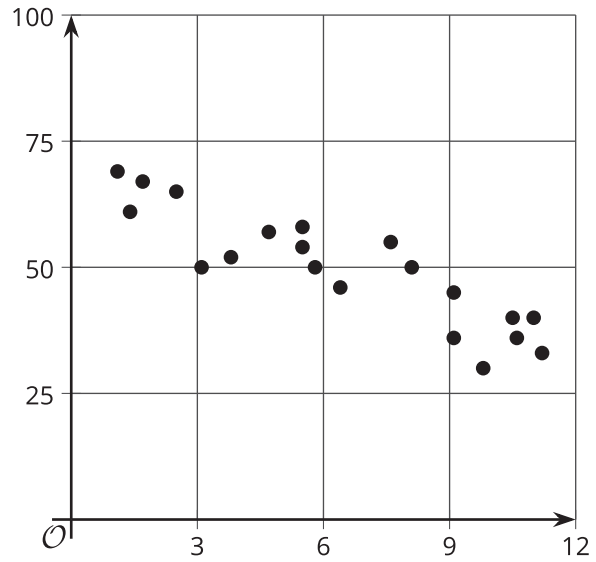
B



C



D



Student Response

Sample responses:

- A, B, and C go together because they have a line that models a connection between the variables.
- A, B, and D go together because the points are fairly close together.
- A, C, and D go together because all the points have a positive y-coordinate.
- B, C, and D go together because the data show a negative trend.

Activity Synthesis

Invite each group to share one reason why a particular set of 3 go together. Record and display the responses for all to see. After each response, ask the class if they agree or disagree. Since there is no single correct answer to the question

of which three go together, attend to students' explanations and ensure the reasons given are correct.

During the discussion, ask students to explain the meaning of any terminology they use, such as "trend," "model," or "variable." Also, press students on unsubstantiated claims.

During the discussion, introduce new vocabulary:

- A **positive association** is a relationship between 2 quantities where one tends to increase as the other increases.
- A **negative association** is a relationship between 2 quantities where one tends to decrease as the other increases.

21.2 Fitting Lines

20 min

Activity Narrative

There is a digital version of this activity.

In this activity, students draw their own linear model to fit the data in a scatter plot. In one scatter plot, the data points are nearly linear, and in another there is much more variation in the data. A discussion follows about what makes some lines a better fit than others (MP3).

In the digital version of the activity, students use an applet to fit a linear model to the data. The applet allows students to drag two points on a line around the graph to find a good linear model. Use the digital version if available to allow students to try different models easily.

Monitor for groups who use these strategies:

- Find the middle point of the data and twist the line to find a slope that fits the rest of the data.
- Make a visual estimate of an appropriate slope, and then slide the line around to go through the middle of the data.

Standards

Addressing **8.SP.A.2**

Instructional Routines

- 5 Practices

Launch

Arrange students in groups of 2. Provide each student a piece of dried pasta and a straightedge.

Tell students that they may use the pasta to try different lines to see what might fit the data best before actually drawing a line with their straightedge and pencil.

Select students with different strategies, such as those described in the *Activity Narrative*, to share later.

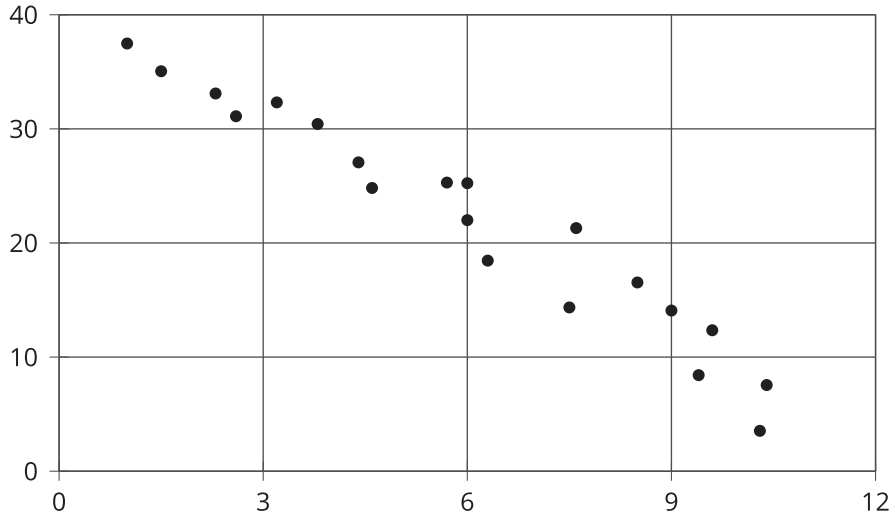
Access for Students with Disabilities

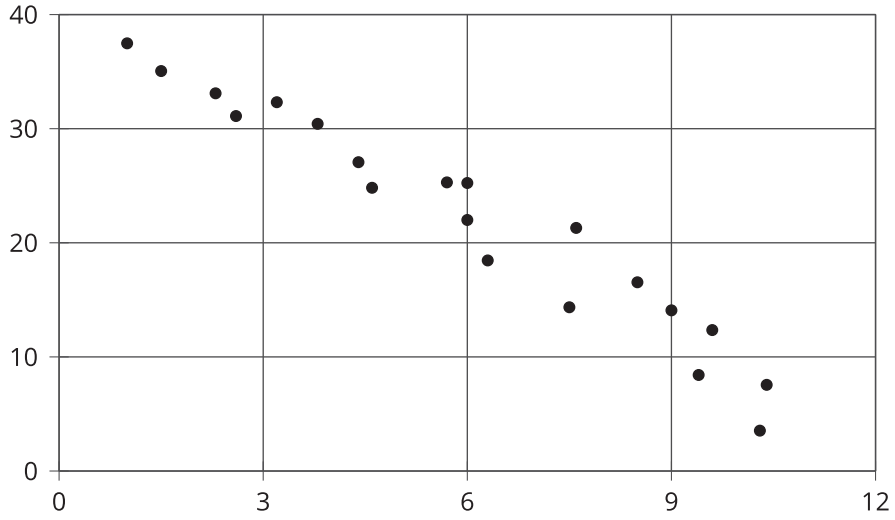
Action and Expression: Develop Expression and Communication. To help get students started, display sentence frames such as "We are trying to . . .," "Let's try . . .," and "I noticed _____ so I . . . ."
Supports accessibility for: Language, Organization

Student Task Statement

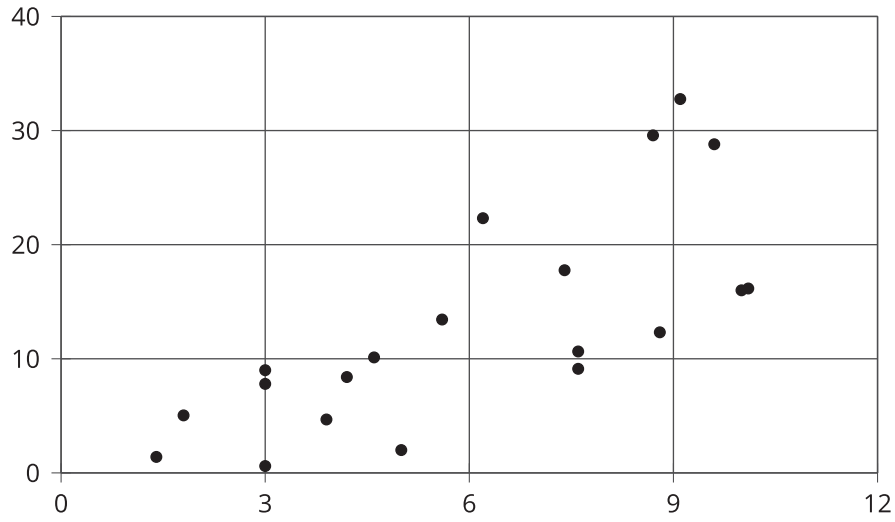
Your teacher will give you a piece of pasta and a straightedge.

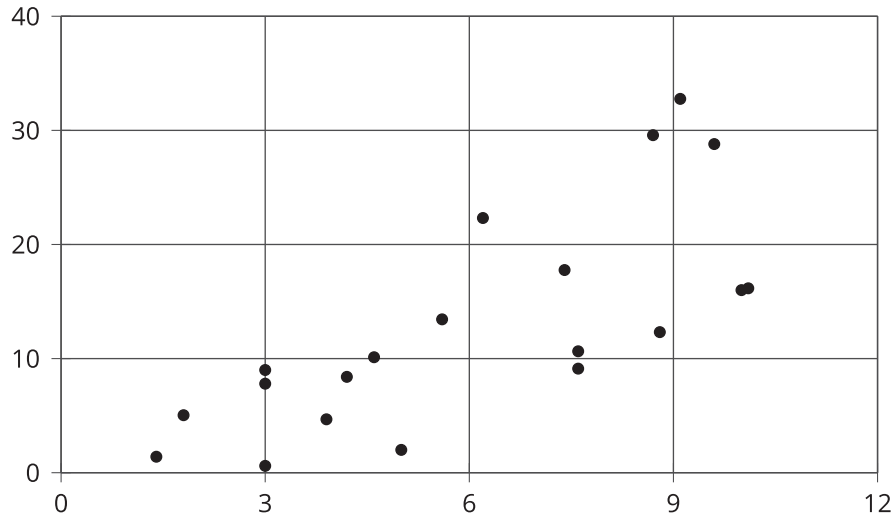
1. Here are two copies of the same scatter plot. Experiment with drawing lines to fit the data. Draw the line that you think best fits the data. Compare it with a partner's.





2. Here are two copies of another scatter plot. Experiment with drawing lines to fit the data. Draw the line that you think best fits the data. Compare it with a partner's.



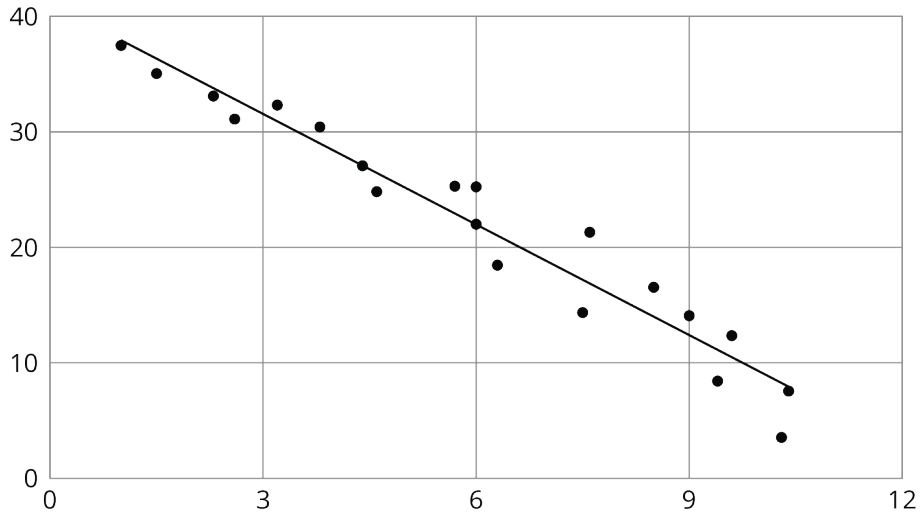


3. In your own words, describe what makes a line fit a data set well.

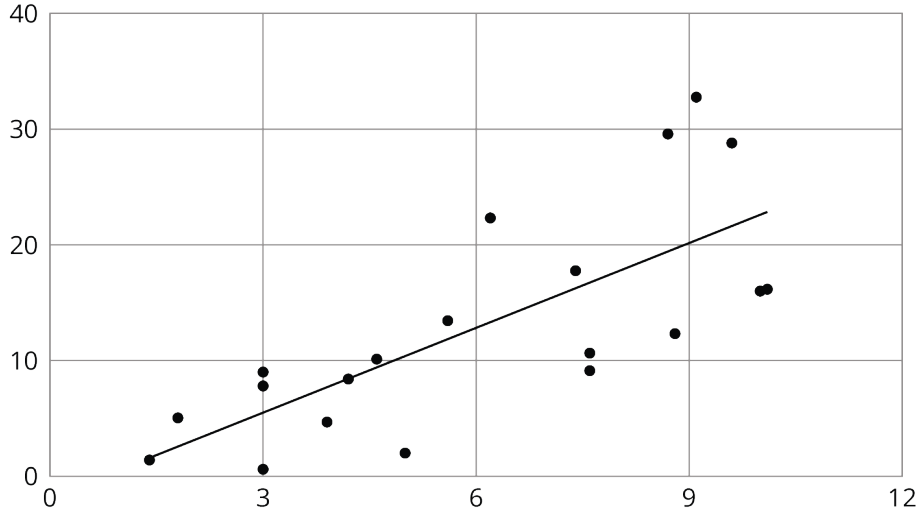
Student Response

Sample responses:

1. An appropriate line created by students will have a negative trend and go through the middle of the cloud of points. For reference, here is a line made using a least-squares regression:



2. An appropriate line created by students will have a positive trend and go through the middle of the cloud of points. For reference, here is a line made using a least-squares regression:



3. ◦ The line passes through the middle of the points.
   ◦ The line has a slope that shows the trend of the points.
   ◦ All of the points are as close to the line as it is possible to get them.

Activity Synthesis

The purpose of this discussion is to look at some strategies for drawing a line that fits the data well.

Invite previously selected groups to share how they found a good linear model. Sequence the discussion of the strategies in the order listed in the *Activity Narrative*. If possible, record and display their work for all to see.

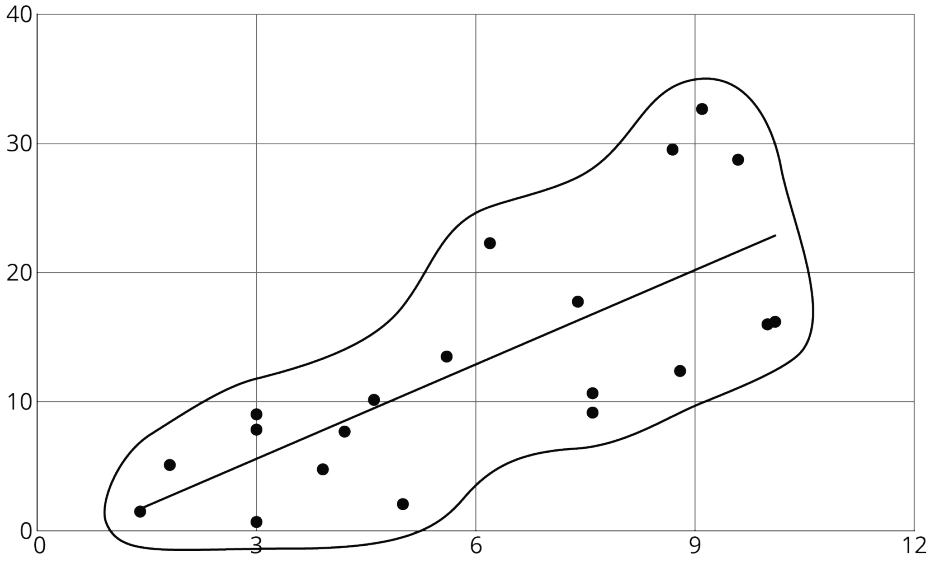
Connect the different responses to the learning goals by asking questions such as:

- "An important part of a line is its slope. How did you find a good slope for a line that can fit the data?" (We twisted the line to be more or less steep to find an angle that kind of goes along with the points.)
- "Another important part of a line is its position on the plane. Why do you think your line fits best where it is rather than sliding it up, down, left, or right?" (With the line where we put it, it goes through the middle of the data so that

there are some points on either side of it. Moving it would not seem as balanced to us.)

If desired and time allows, demonstrate this procedure:

1. Enclose all of the points in the scatter plot with a blob.
2. Use a straightedge to draw a line "through the middle" of the blob. Some students find it helpful to think of the blob as a hot dog bun and the line as the hot dog.



21.3 Good Fit Bad Fit

Optional

15 min

Activity Narrative

If students understand what makes for a good fit from the previous activity, then this activity may be considered optional. The next activity will give students less scaffolded practice deciding if a line fits the data well.

Students have seen linear models for data in a previous lesson. In this activity, students begin to determine what makes a good model for data. They compare two different lines with the same data set to determine which model fits the data better. A formal, quantitative discussion of lines that best fit data will come in later grades. At this stage, students are only asked to informally determine whether the line fits the data well based on how close the points are to the line.

Standards

Addressing **8.SP.A.2**

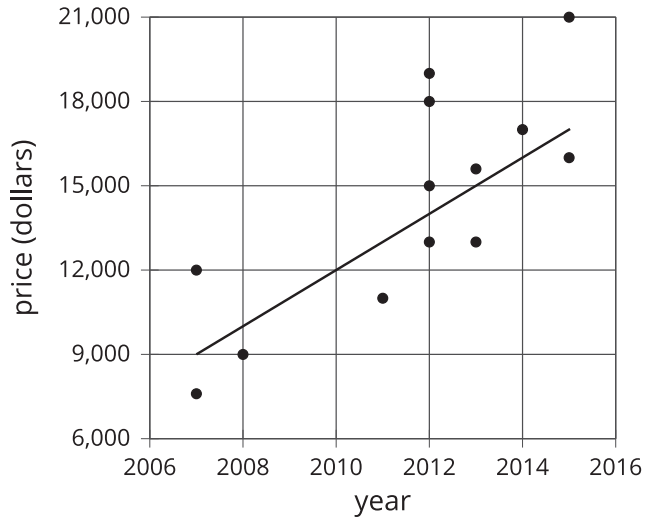
Instructional Routines

- MLR1: Stronger and Clearer Each Time
- Notice and Wonder

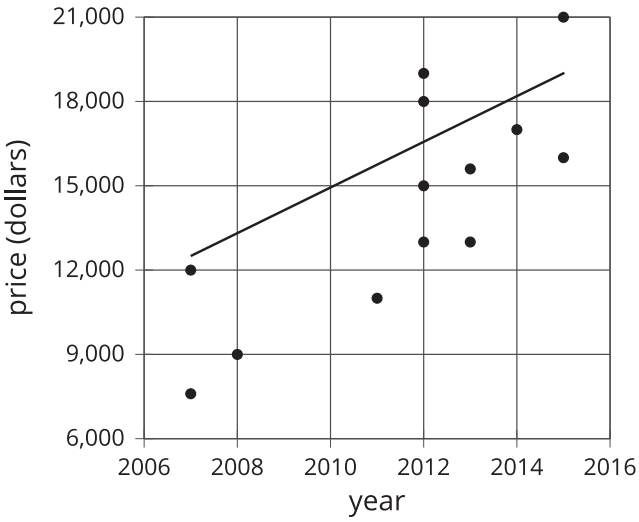
Launch

Display the scatter plot for all to see and ask students, "What do you notice? What do you wonder?"

A



B



Students might notice:

- There are 17 points plotted.
- The scatter plot shows a positive association (or a positive trend).
- The horizontal axis represents years, and the vertical axis represents price.

Students might wonder:

- What are these the prices of?
- Why do later years have a higher price?

Tell students that these are all prices of used cars that are all the same make and model that are for sale. For each car, the scatter plot shows its year of manufacture and the price at which it is being sold. Ask students a few questions to familiarize themselves with the graph, like:

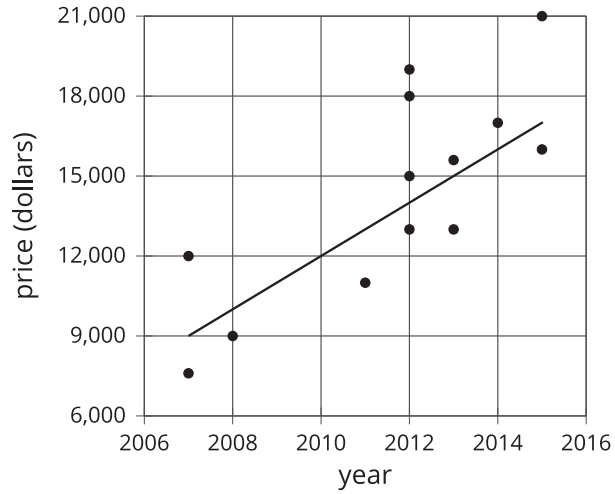
- "What do we mean when we say 'used car'?" (It is not a new car. It has already been owned by someone.)
- "How many of these cars were made in 2012?" (4)
- "What is the price of the car made in 2008?" ($9,000)
- "The data show a positive association. What does that mean in this situation?" (Cars made in a later year tend to have a higher price.)

Tell students that in this task, they are going to see 2 different models for this set of data.

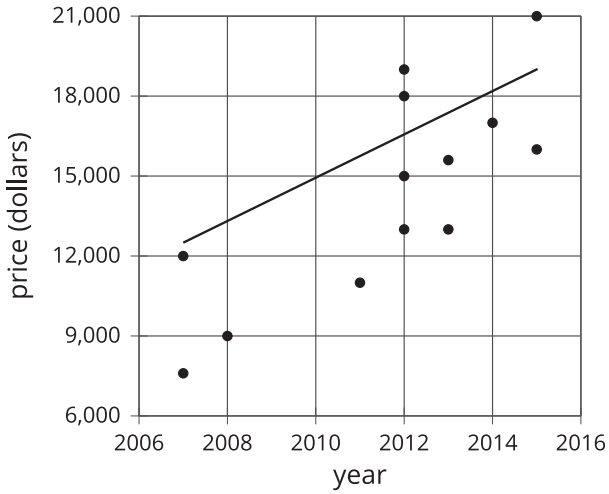
Student Task Statement

The scatter plots both show the year and price for the same 13 used cars. However, each scatter plot shows a different model for the relationship between year and price.

A



B



1. Look at Diagram A.
   a. For how many cars does the model in Diagram A make a good prediction of its price?
   b. For how many cars does the model underestimate the price?
   c. For how many cars does it overestimate the price?
2. Look at Diagram B.
   a. For how many cars does the model in Diagram B make a good prediction of its price?
   b. For how many cars does the model underestimate the price?
   c. For how many cars does it overestimate the price?
3. For how many cars does the prediction made by the model in Diagram A differ by more than $3,000? What about the model in Diagram B?
4. Which model does a better job of predicting the price of a used car from its year?

Student Response

Sample responses:

1. Nine points are fairly close to the line, so 9 prices are well predicted. The model underestimates the price of 7 cars and overestimates the price of 6 cars.
2. Only the point $(2007, 12,000)$ is close to the prediction. The model underestimates the price of 3 cars and overestimates the price of 10 cars.
3. Model A over- or underestimates the price of 4 cars by about $3,000 or more. For model B, there are 6 cars whose price differs from the model by more than $3,000.
4. Model A does a better job of predicting the prices of these cars based on their year.

Activity Synthesis

The purpose of this discussion is for students to see some strategies for evaluating the fit of a model.

Some questions for discussion:

- "Which model did a better job of fitting with the data?" (Model A)

- "What were some things that helped you determine which model was better for this data?" (The line went through the "middle" of the data. There were some points on each side of the line, so that it looks to be in the middle. The model predicts the values fairly well for most points.)
- "If a person was looking to buy a used car made in 2006 and incorrectly used Model B, approximately how much money would they be predicting to pay? If they used Model A?" (Model B predicts the cost to be about $12,000 instead of $8,000 from Model A, which is a $4,000 difference, or a 50% increase in price!)

We say that Model A "fits the data" better than Model B, or that model A is a "better fit."

Access for English Language Learners

MLR1 Stronger and Clearer Each Time. Before the whole-class discussion, give students time to meet with 2–3 partners to share and get feedback on their first draft response to the last question. Invite listeners to ask questions and give feedback that will help their partner clarify and strengthen their ideas and writing. Give students 3–5 minutes to revise their first draft based on the feedback they receive.
Advances: Writing, Speaking, Listening

21.4 Practice Fitting Lines

10 min

Activity Narrative

This activity gives students additional practice finding linear models that match the association of the data. In the first scatter plot, students are given a linear model that has a good slope, but is shifted up from the center of the data. In the second set of data, students are given a linear model that goes through the middle of the data, but has a slope that is too steep. Students are given the opportunity to correct these issues by drawing their own linear models on the same scatter plots.

Standards

Addressing **8.SP.A.2**

Instructional Routines

- MLR2: Collect and Display

Launch

Arrange students in groups of 2. Give students 2 minutes quiet work time followed by partner discussion and whole-class discussion.

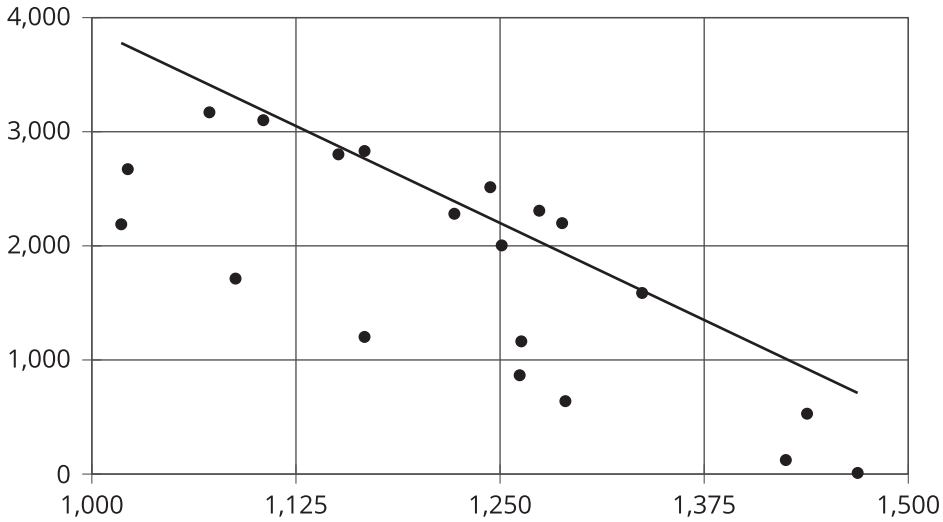
Ask students why a line might be added to a scatter plot (to help predict additional values, to show a positive or negative association). Tell students that they will have a chance to practice adding lines to scatter plots by first critiquing a given line and then improving the linear model by drawing their own line for the same scatter plot.

Access for English Language Learners

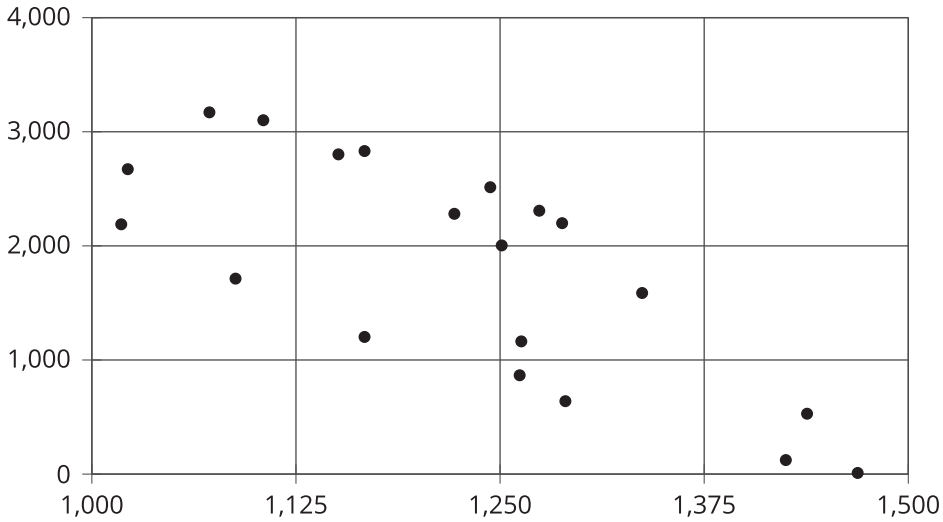
MLR2 Collect and Display. Collect the language students use to describe the linear models they draw. Display words and phrases such as "through the middle," "points above and below," or "slope." During the synthesis, invite students to suggest ways to update the display: "What are some other words or phrases we should include?" Invite students to borrow language from the display as needed.
Advances: Conversing, Reading

Student Task Statement

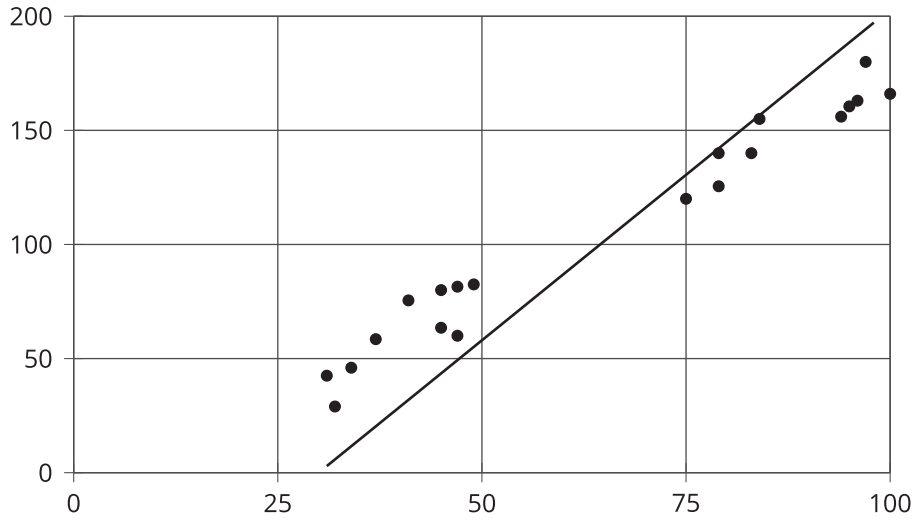
1. Is this line a good fit for the data? Explain your reasoning.



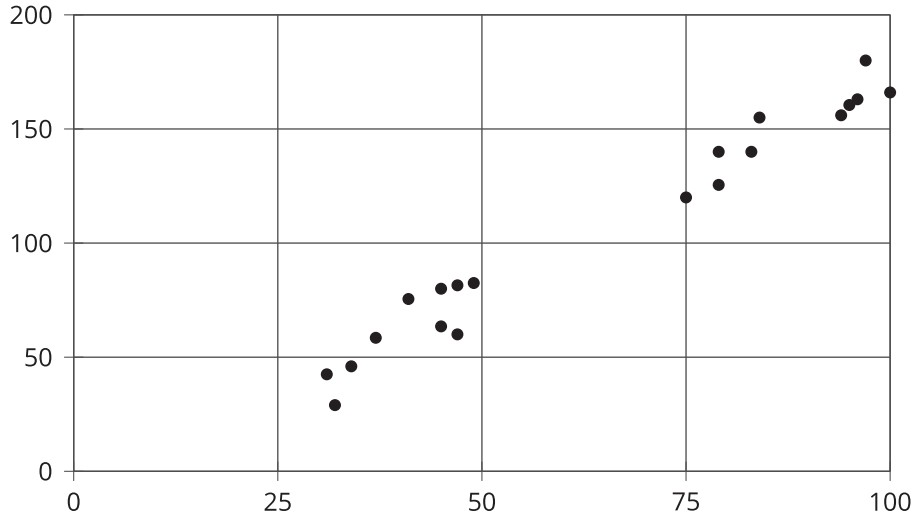
2. Draw a line that fits the data better.



3. Is this line a good fit for the data? Explain your reasoning.



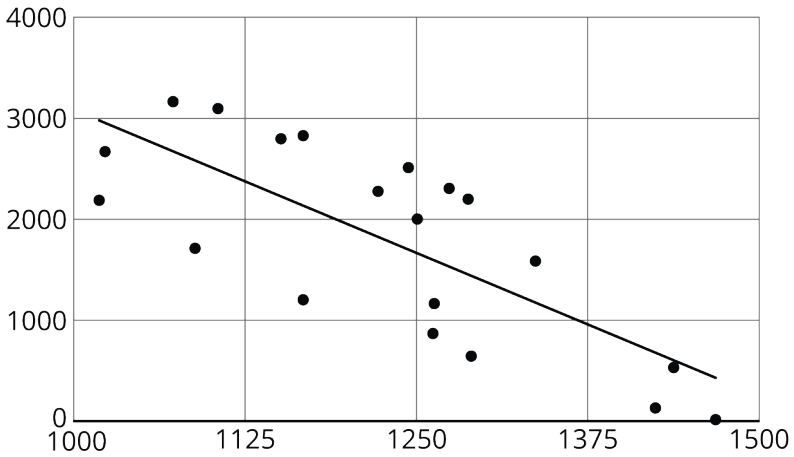
4. Draw a line that fits the data better.



Student Response

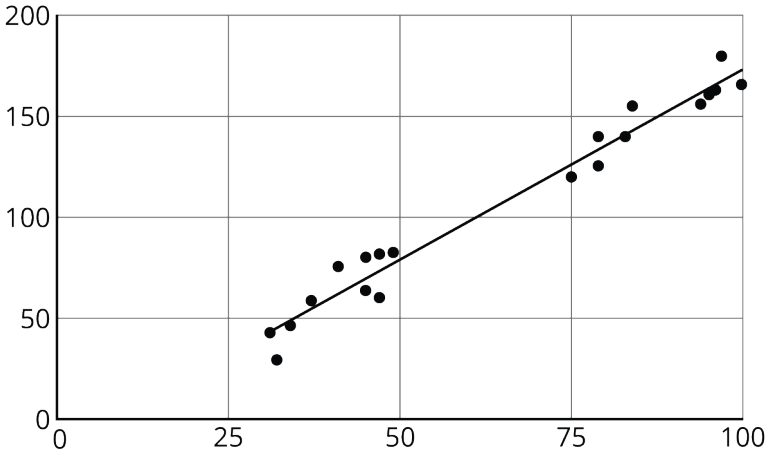
1. No. There are many more points below the line than above it. A line with a good fit should be below this drawn line.

2. Sample response:

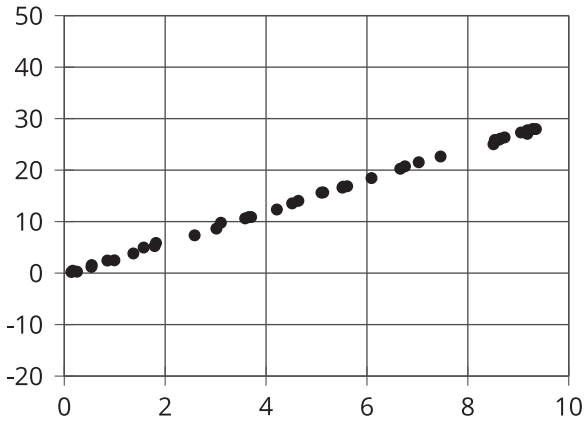


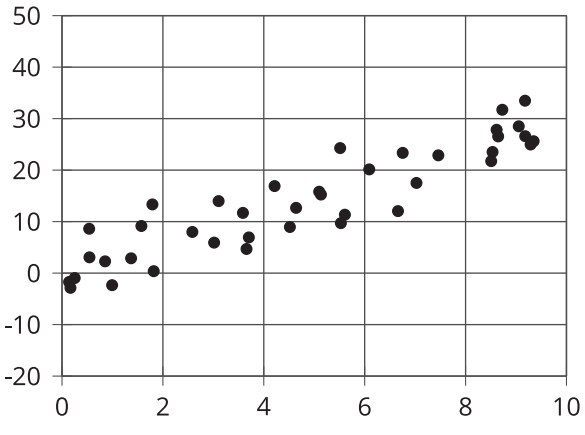
3. Not great. The line could be tilted to have a slightly smaller slope to fit the line better.

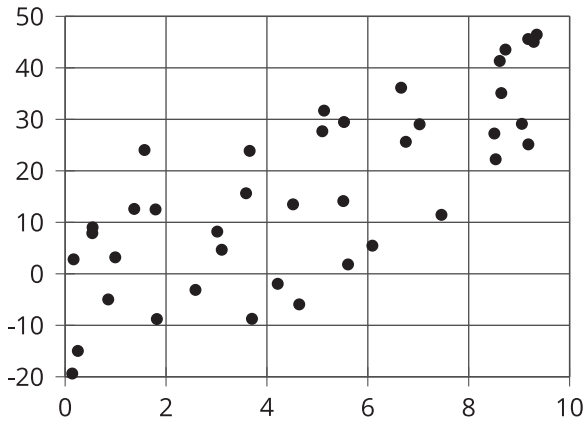
4. Sample response:



Are You Ready for More?







These scatter plots were created by multiplying the x-coordinate by 3 then adding a random number between two values to get the y-coordinate. The first scatter plot added a random number between -0.5 and 0.5 to the y-coordinate. The second scatter plot added a random number between -8 and 8 to the y-coordinate. The third scatter plot added a random number between -20 and 20 to the y-coordinate.

1. For each scatter plot, draw a line that fits the data.
2. Explain why some were easier to do than others.

Extension Student Response

Sample response:

1. The line $y = 3x$ (or equivalent) would be a good fit for each graph.
2. Because the points are closer together in the first 2 graphs, it is easier to draw the line for those. Although there is some increasing trend in the third graph, it is difficult to draw a good line because the points are so spread out.

Activity Synthesis

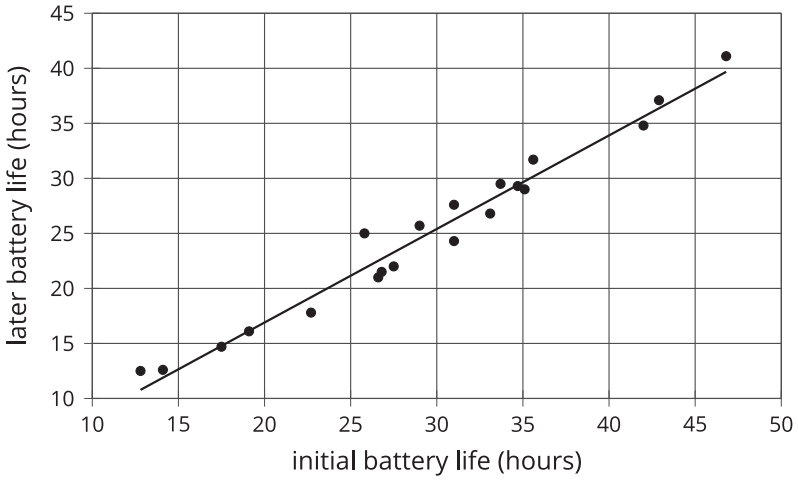
The purpose of the discussion is for students to recognize the important aspects of a linear model for a set of data.

Consider asking some of the following questions.

- "Compare the given lines to the ones you drew in terms of their slopes and vertical positions." (In the first scatter plot, they should have about the same slope, but the given line is shifted up. In the second scatter plot, they should both go through a point near the center of the data, but have different slopes.)
- "What are some ways you thought about trying to find a good line to draw?"

Lesson Synthesis

Display the scatter plot for all to see.



To highlight the main ideas from today's lesson about associations and trend lines, ask:

- "How would you describe the relationship between the later battery life and initial battery life?" (There is a positive association. Or, as the initial battery life increases, the later battery life tends to increase.)
- "How can we tell if a line is a good fit for the data in a scatter plot?" (It goes through the middle of the data, the sign of the slope matches the sign of the association, or the points are as close as possible to the line.)

21.5 This Is One Way to Do It

Cool-down

5 min

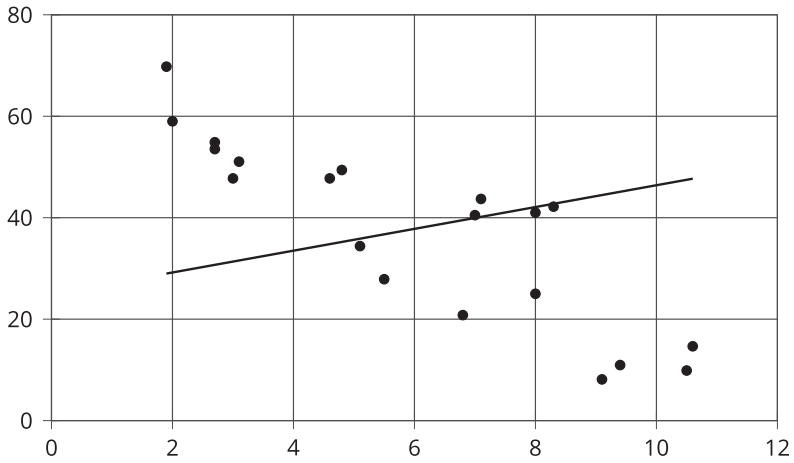
This activity addresses some misconceptions students may have while thinking about linear models. Students critique attempts to draw a line that fits data in a scatter plot. In particular, two lines are drawn that resolve some concerns when fitting lines but disregard others.

Standards

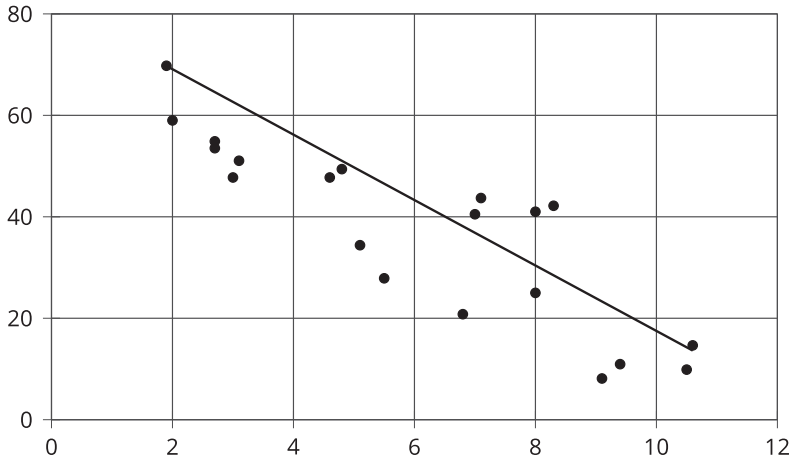
Addressing **8.SP.A.2**

Student Task Statement

1. Elena said, "I think this line is a good fit because half of the points are on one side of the line and half of the points are on the other side." Do you agree? Explain your reasoning.



2. Noah said, "I think this line is a good fit because it passes through the leftmost point and the rightmost point." Do you agree? Explain your reasoning.



Student Response

1. Disagree. Sample response: The line is not a good fit because the data show a negative association, but the line has a positive slope.
2. Disagree. Sample responses: The line is not a good fit because most of the points are below it and the trend of the scatter plot is steeper than the slope of the graph.

Responding to Student Thinking

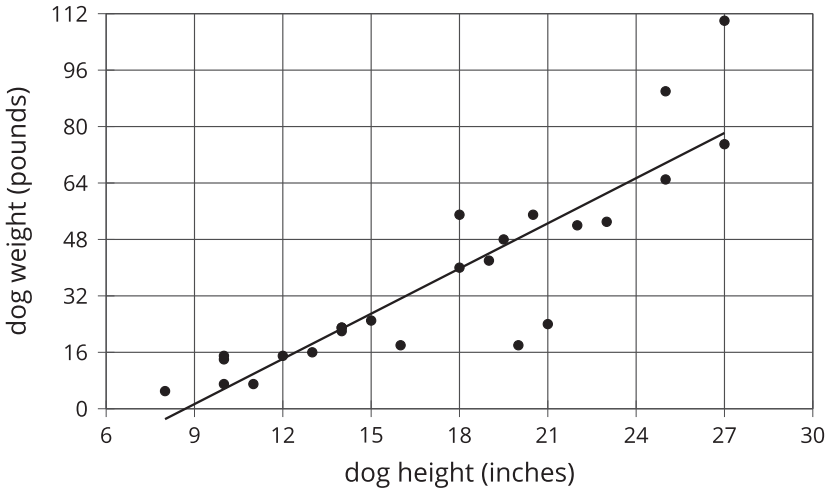
Points to Emphasize

If students struggle with analyzing lines of fit, plan to focus on this as opportunities arise over the next several lessons. For example, in the activity referred to here, invite multiple students to share their thinking about the associations between independent and dependent variables.

Lesson 21 Summary

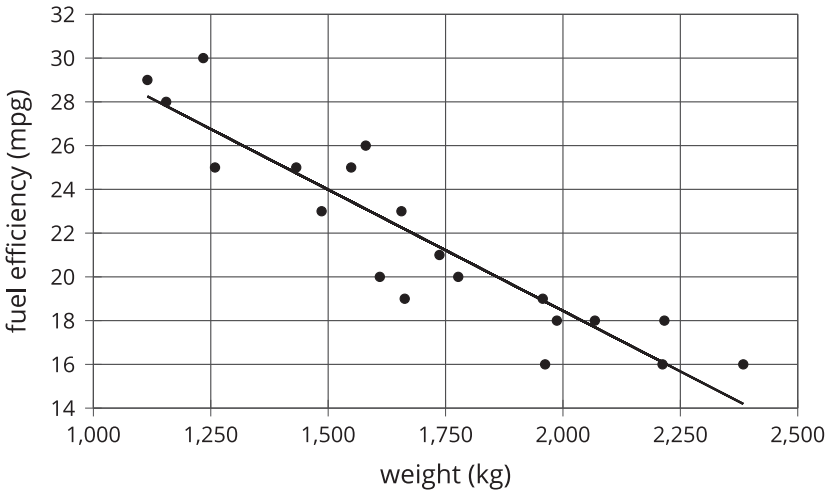
When a linear function fits data well, we say there is a "linear association" between the variables. For example, the relationship between height and weight for 25 dogs with the linear function whose graph is shown in the scatter plot.

We say there is a **positive association** between dog height and dog weight because knowledge about one variable helps predict the other variable, and when one variable increases, the other tends to increase as well.



What do you think the association between the weight of a car and its fuel efficiency is?

We say that there is a **negative association** between fuel efficiency and weight of a car because knowledge about one variable helps predict the other variable, and when one variable increases, the other tends to decrease.



Glossary

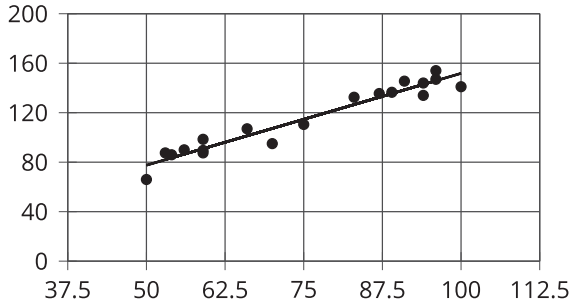
- negative association
- positive association

Lesson 21 Practice Problems

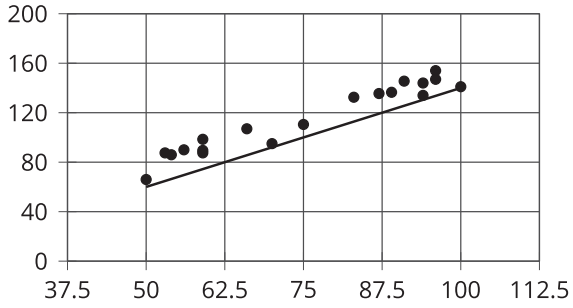
1 Student Task Statement

For the same data, two different models are graphed. Which model more closely matches the data? Explain your reasoning.

A



B



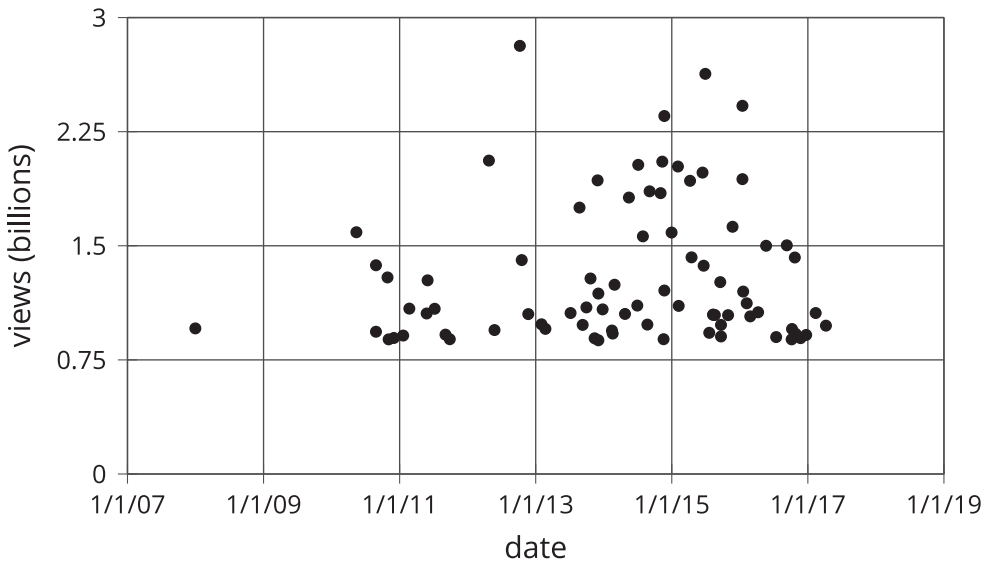
Solution

Sample response: Model A more closely matches the data. In Model B, most of the points are above the line in the graph. In Model A, the points are more evenly arranged around the line.

2 from Unit 5, Lesson 19

Student Task Statement

Here is a scatter plot that shows the most popular videos in a 10-year span.



a. Use the scatter plot to estimate the number of views for the most popular video in this 10-year span.

b. Estimate when the 4th most popular video was released.

Solution

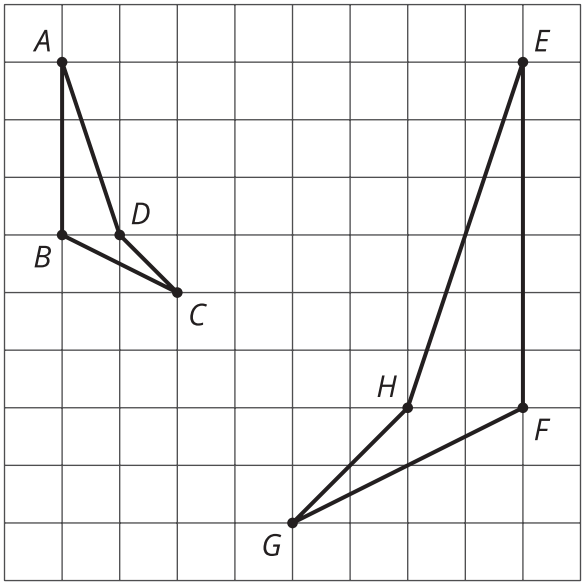
a. The most popular video has roughly 2.8 billion views.

b. Late 2014

3 from Unit 2, Lesson 12

Student Task Statement

Show that the two figures are similar by identifying a sequence of translations, rotations, reflections, and dilations that takes one figure to the other.



Solution

Sample response: Translate H to D, reflect across a vertical line through D, and then dilate using a scale factor of $\frac{1}{2}$ centered at D.

4 from Unit 4, Lesson 2

Student Task Statement

Select **all** the expressions that are equivalent to $-36x + 54y - 90$.

A. $-9(4x - 6y - 10)$

B. $-18(2x - 3y + 5)$

C. $-6(6x + 9y - 15)$

D. $18(-2x + 3y - 5)$

E. $-2(18x - 27y + 45)$

F. $2(-18x + 54y - 90)$

Solution

B, D, E