Lesson 8

Analyzing Bivariate Data

Goals

- Create a scatter plot and draw a line to fit bivariate data, and identify (orally and in writing) outliers that appear in the data.
- Interpret (orally and in writing) features of a scatter plot with a line of fit, including outliers, slope of the line, and clustering.

Learning Target

I can analyze a set of data to determine associations between two variables.

Lesson Narrative

In this lesson, students bring together everything they have studied in the unit so far to analyze and interpret bivariate data in context. They create a scatter plot, identify outliers, fit a line, then determine and interpret the slope of the line. They compare actual and predicted values using a linear model. Students also analyze the data collected in an earlier lesson about their classmates.

Student Learning Goal

Let's analyze data like a pro.

Access for Students with Diverse Abilities

- Action and Expression (Activity 2)

Access for Multilingual Learners

- MLR1: Stronger and Clearer Each Time (Activity 2)
- MLR8: Discussion Supports (Activity 1)

Required Materials

Materials to Gather

- Dried linguine pasta: Activity 1
- Straightedges: Activity 1

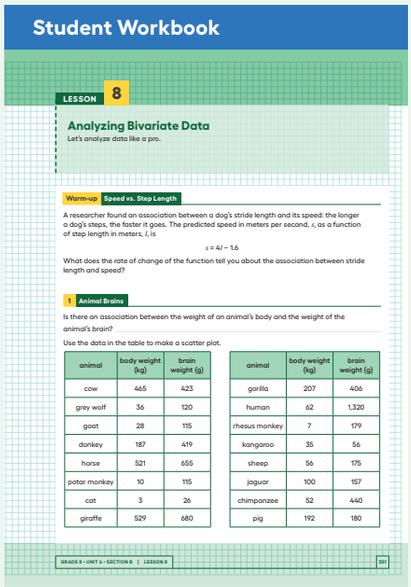
Activity 1:
For the digital version of the activity, acquire devices that can run the applet.

Activity 2:
Class data from an earlier lesson in the unit on number of pages and weight of books.

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

Lesson Timeline

Warm-up	Activity 1	Activity 2	Lesson Synthesis	Assessment
5 min	15 min	15 min	10 min	5 min
				Cool-down



Student Workbook

LESSON 8

Analyzing Bivariate Data

Let's analyze data like a pro.

Warm-up Speed vs. Step Length

A researcher found an association between a dog's stride length and its speed: the longer a dog's steps, the faster it goes. The predicted speed in meters per second, s, as a function of step length in meters, l, is

$s = 4l - 1.6$

What does the rate of change of the function tell you about the association between stride length and speed?

1 Animal Brains

Is there an association between the weight of an animal's body and the weight of the animal's brain?

Use the data in the table to make a scatter plot.

animal	body weight (kg)	brain weight (g)
cow	465	423
grey wolf	36	120
goat	28	115
donkey	187	419
horse	521	655
potar monkey	10	115
cat	3	26
giraffe	529	680

animal	body weight (kg)	brain weight (g)
gorilla	207	406
human	62	1,320
rhesus monkey	7	179
kangaroo	35	56
sheep	56	175
jaguar	100	157
chimpanzee	52	440
pig	192	180

GRADE 8 • UNIT 6 • SECTION B | LESSON 8 231

Warm-up

Speed vs. Step Length

5 min

Activity Narrative

The purpose of this *Warm-up* is for students to interpret the rate of change of a function to describe a trend. It is used to help students recall the role of slope in showing associations for data that can be fit with a linear model in anticipation of looking for associations in this lesson.

Launch

Give students 2 minutes of quiet work time followed by a whole-class discussion. Remind students that the phrase "rate of change" is the interpretation of the slope of a line for the situation.

Student Task Statement

A researcher found an association between a dog's stride length and its speed: the longer a dog's steps, the faster it goes. The predicted speed in meters per second, s, as a function of step length in meters, l, is

$$s = 4l - 1.6$$

What does the rate of change of the function tell you about the association between stride length and speed?

For every 1 meter increase in a dog's step length, its speed increases by 4 meters per second.

Activity Synthesis

Ask students to share what the rate of change of this function tells them about the trend. Record and display their responses for all to see. To include more students in the conversation, consider asking some of the following questions:

- *"Does anyone agree or disagree with this reasoning? Why?"*
- *"Did anyone reason about the rate of change in a different way?"*
- *"Did anyone reason about the rate of change in the same way but would describe the trend differently?"*
- *"Does anyone want to add on to ____'s reasoning?"*

Activity 1
Animal Brains

15 min

Activity Narrative

There is a digital version of this activity.

All of the information from this section about scatter plots comes into play as students analyze data about animal body and brain weights. Students begin with a table of data and create a scatter plot. After seeing the scatter plot, students pick out any outliers and fit a line to the scatter plot. Finally, the slope of the line is estimated and its meaning interpreted in context.

In the digital version of the activity, students use an applet to analyze and 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.

Launch

Arrange students in groups of 2. Give students 5 minutes of quiet work time followed by 5 minutes of partner discussion and 5 minutes of whole-class discussion.

Student Task Statement

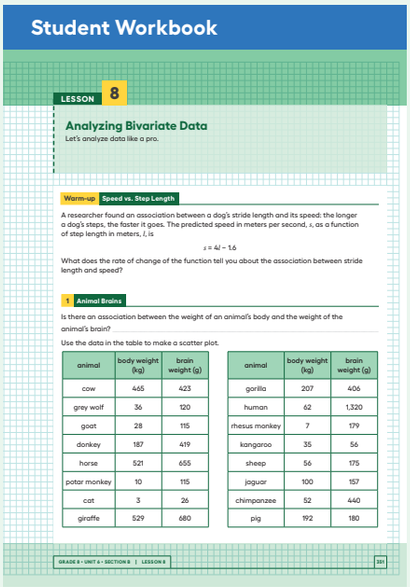
Is there an association between the weight of an animal's body and the weight of the animal's brain?

Use the data in the table to make a scatter plot. Are there any outliers?

animal	body weight (kg)	brain weight (g)
cow	465	423
grey wolf	36	120
goat	28	115
donkey	187	419
horse	521	655
potar monkey	10	115
cat	3	26
giraffe	529	680

animal	body weight (kg)	brain weight (g)
gorilla	207	406
human	62	1,320
rhesus monkey	7	179
kangaroo	35	56
sheep	56	175
jaguar	100	157
chimpanzee	52	440
pig	192	180

Student Workbook



LESSON 8

Analyzing Bivariate Data

Let's analyze data like a pro.

Warm-up Speed vs. Step Length

A researcher found an association between a dog's stride length and its speed: the longer a dog's steps, the faster it goes. The predicted speed in meters per second, s, as a function of step length in meters, l, is

$s = 4l - 1.6$

What does the rate of change of the function tell you about the association between stride length and speed?

1 Animal Brains

Is there an association between the weight of an animal's body and the weight of the animal's brain?

Use the data in the table to make a scatter plot.

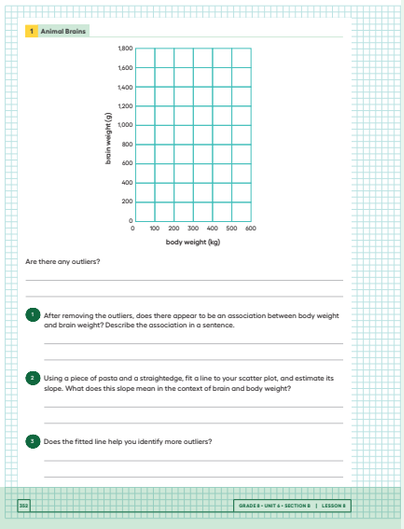
animal	body weight (kg)	brain weight (g)
cow	465	423
grey wolf	36	120
goat	28	115
donkey	187	419
horse	521	655
potar monkey	10	115
cat	3	26
giraffe	529	680

animal	body weight (kg)	brain weight (g)
gorilla	207	406
human	62	1,320
rhesus monkey	7	179
kangaroo	35	56
sheep	56	175
jaguar	100	157
chimpanzee	52	440
pig	192	180

Building on Student Thinking

When estimating slope, some students won't use the scales of the axes correctly, so the slope is reported incorrectly. Some students may not notice the different units of weight used on each axis.

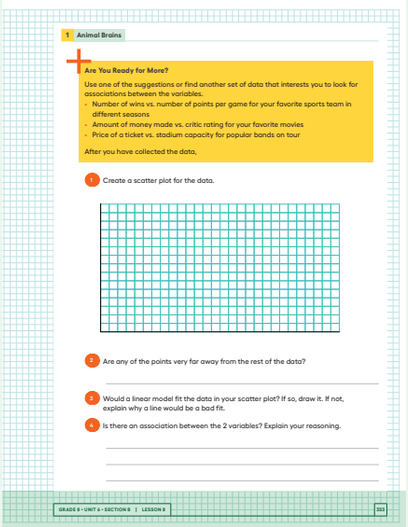
Student Workbook


1 Animal Brains

Are there any outliers?

1. After removing the outliers, does there appear to be an association between body weight and brain weight? Describe the association in a sentence.
2. Using a piece of pasta and a straightedge, fit a line to your scatter plot, and estimate its slope. What does this slope mean in the context of brain and body weight?
3. Does the fitted line help you identify more outliers?

Student Workbook


1 Animal Brains

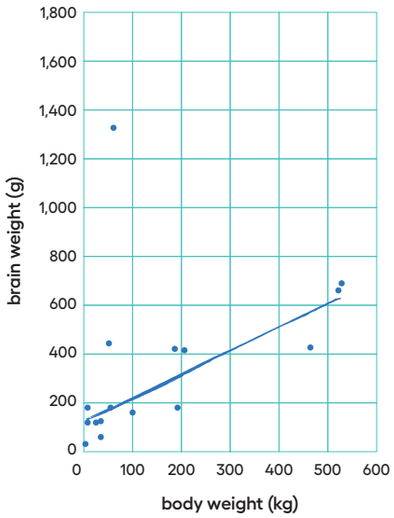
Are You Ready for More?

Use one of the suggestions or find another set of data that interests you to look for associations between the variables.

- Number of wins vs. number of points per game for your favorite sports team in different seasons
- Amount of money made vs. critic rating for your favorite movies
- Price of a ticket vs. stadium capacity for popular bands on tour

After you have collected the data,

1. Create a scatter plot for the data.
2. Are any of the points very far away from the rest of the data?
3. Would a linear model fit the data in your scatter plot? If so, draw it. If not, explain why a line would be a bad fit.
4. Is there an association between the 2 variables? Explain your reasoning.



The human data is an outlier.

1. After removing the outliers, does there appear to be an association between body weight and brain weight? Describe the association in a sentence.

   There seems to be a positive, linear association between body and brain weights.

2. Using a piece of pasta and a straightedge, fit a line to your scatter plot, and estimate its slope. What does this slope mean in the context of brain and body weight?

   The slope is about 1. This means that for every increase of 1 kilogram of body weight, the model predicts an increase of 1 gram of brain weight.

3. Does the fitted line help you identify more outliers?

   Sample response: Yes, the data for the chimpanzee is also far from the line and may be an outlier.

Are You Ready for More?

Use one of the suggestions or find another set of data that interests you to look for associations between the variables.

- Number of wins vs. number of points per game for your favorite sports team in different seasons
- Amount of money made vs. critic rating for your favorite movies
- Price of a ticket vs. stadium capacity for popular bands on tour

After you have collected the data,

1. Create a scatter plot for the data.
2. Are any of the points very far away from the rest of the data?
3. Would a linear model fit the data in your scatter plot? If so, draw it. If not, explain why a line would be a bad fit.
4. Is there an association between the 2 variables? Explain your reasoning.

   Answers vary

Activity Synthesis

The goal of this discussion is to ensure students can make sense of the data given all the tools from this unit.

Consider asking some of the following questions:

- *"Which data did you consider outliers?"*

  human and chimpanzee

- *"How did you determine your fitted line?"*

  I tried to get some points on either side of the line so that it would go through the middle of the points.

- *"Let's assume the trend you found continues past the end of the scatter plot. A Tyrannosaurus rex is a dinosaur that is estimated to have a body weight of about 8,000 kg. What do you expect its brain weight to be?"*

  About 8,000 g or 8 kg

Based on the fossils of Tyrannosaurus rex and what they understand about the dinosaur's anatomy, scientists believe that their brains would've weighed about 1 kilogram. The trend does not appear to continue, at least not in a linear way.

Access for Multilingual Learners (Activity 1, Synthesis)

MLR8: Discussion Supports. For each observation that is shared, invite students to turn to a partner and restate what they heard using precise mathematical language. *Advances: Listening, Speaking*

Activity 2

Average Page Weight

15 min

Activity Narrative

There is a digital version of this activity.

In this activity students create another scatter plot to analyze the data they collected about books in a previous lesson. A suggested linear model is compared to the data and a particular point is identified in both the scatter plot and data table.

Although the scatter plots are left to students to organize, the only linear model considered is $y = \frac{1}{50}x$, where y represents the weight in ounces and x represents the number of pages. If possible, identify any groups who have axes switched to bring up in the discussion.

In the digital version of the activity, students use an applet to create a scatter plot and draw the linear model. The applet allows students to input the data into a table and create a scatter plot. Use the digital version if available to allow students to focus on analyzing the scatter plot without worrying about construction.

Access for Students with Diverse Abilities (Activity 2, Launch)

Action and Expression: Provide Access for Physical Action.
Provide access to tools and assistive technologies such as a graphing calculator, graphing software, or the digital version of this activity.
Supports accessibility for: Visual-Spatial Processing, Conceptual Processing, Organization

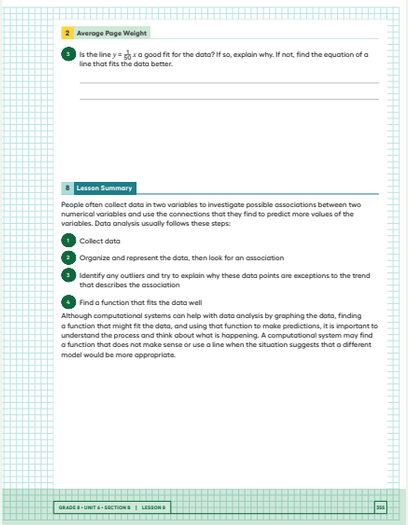
Student Workbook


2 Average Page Weight

Earlier, your class gathered data on the number of pages and weight of books.

1. Sometimes the average weight per page of a book is about 1 ounce per 50 pages. Is this true for any of the books you measured? _____
2. Make a scatter plot for the number of pages and weight data, and describe any association.

Student Workbook


2 Average Page Weight

3. Is the line $y = \frac{1}{50}x$ a good fit for the data? If so, explain why. If not, find the equation of a line that fits the data better.

8 Lesson Summary

People often collect data in two variables to investigate possible associations between two numerical variables and use the connections that they find to predict more values of the variables. Data analysis usually follows these steps:

1. Collect data
2. Organize and represent the data, then look for an association
3. Identify any outliers and try to explain why these data points are exceptions to the trend that describes the association
4. Find a function that fits the data well

Although computational systems can help with data analysis by graphing the data, finding a function that might fit the data, and using that function to make predictions, it is important to understand the process and think about what is happening. A computational system may find a function that does not make sense or use a line when the situation suggests that a different model would be more appropriate.

Launch

Arrange students in groups of 2. Give students 10 minutes of quiet work time followed by 5 minutes of whole-class discussion.

Display and share the class data collected earlier in the unit.

Student Task Statement

Earlier, your class gathered data on the number of pages and weight of books.

Answers vary based on class data.

1. Sometimes the average weight per page of a book is about 1 ounce per 50 pages. Is this true for any of the books you measured?

   Any data points on the line $y = \frac{1}{50}x$ represent books like this.

2. Make a scatter plot for the number of pages and weight data, and describe any association.

   Graphs should have one axis labeled with weight (oz) and one with number of pages.

   There is a positive, linear association between weight and number of pages, because as the number of pages increases, weight tends to increase.

3. Is the line $y = \frac{1}{50}x$ a good fit for the data? If so, explain why. If not, find the equation of a line that fits the data better.

   The line $y = \frac{1}{50}x$ may be a good fit, because in general, this works for paperback books.

Activity Synthesis

The goal of this activity is for students to use the methods they have learned in this unit to explore data they have collected. To highlight some of the main points, select 2–3 students to respond to each question:

- *"What is the slope of the line that best fits the data?"*

  Answers vary based on class data, but should be close to $\frac{1}{50}$.

If any groups had axes switched, select these groups to show their scatter plots.

- *"How can the the line $y = \frac{1}{50}x$ be changed if the axes are different?"*

  We could switch the x and y then divide by $\frac{1}{50}$ on each side to get $y = 50x$.

- *"What does it mean about a books that is represented by a point is above the line $y = \frac{1}{50}x$? What does it mean about a book represented by a point below the line?"*

  Depending on the axes, a point above the line represents a book with pages that weigh more than $\frac{1}{50}$ ounces on average. Below the line, the books have pages that weigh less than $\frac{1}{50}$ ounces on average.

- *"Suppose you measure pages and weight for a lot of books in the library. When you make a scatter plot, you notice 2 groups of data the seem to line up like parallel lines. What could this mean? Why do you think there may be these clusters?"*

  It may mean there are hardback books and paperback books. Because the weight of the cover is included, that could bring up the average for that type of book.

Lesson Synthesis

The goal of this discussion is to help students reflect on all of the things they have learned about bivariate data in this unit. Consider asking some of the following questions:

- *"What does a point in a scatter plot tell you?"*

  2 measurements about an individual in a population

- *"What is an association between variables?"*

  A trend that suggests that as one variable increases, the other variable tends to increase if it is a positive association or decrease if it is a negative association.

- *"What does a fitted line tell you about the data?"*

  It represents a model that can be used to make predictions about the dependent variable based on the value of the independent variable.

- *"What does the slope of a fitted line tell you about the data?"*

  The amount the dependent variable will increase (or decrease) for a one-unit increase in the independent variable.

Access for Multilingual Learners (Activity 2, Synthesis)

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 "Is the line $y = x$ a good fit for the data? If so, explain why. If not, find the equation of a better line." 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

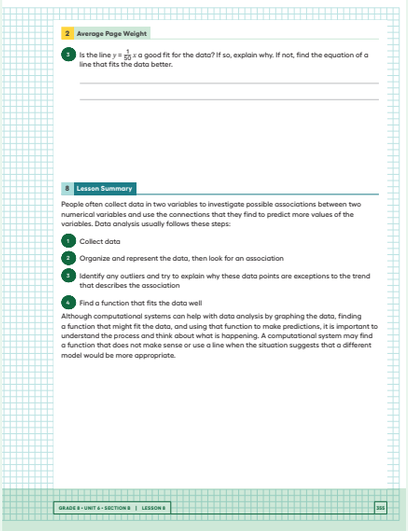
Responding To Student Thinking

Press Pause

By this point in the unit, there should be some student mastery of drawing a linear model and using it to make a prediction. If most students struggle, make time to examine related work in the section referred to here. The Course Guide provides additional ideas for revisiting earlier work.

Unit 6, Section B Associations in Numerical Data

Student Workbook



2 Average Page Weight

3 Is the line $y = \frac{1}{50}x$ a good fit for the data? If so, explain why. If not, find the equation of a line that fits the data better.

8 Lesson Summary

People often collect data in two variables to investigate possible associations between two numerical variables and use the connections that they find to predict more values of the variables. Data analysis usually follows these steps:

1. Collect data
2. Organize and represent the data, then look for an association
3. Identify any outliers and try to explain why these data points are exceptions to the trend that describes the association
4. Find a function that fits the data well

Although computational systems can help with data analysis by graphing the data, finding a function that might fit the data, and using that function to make predictions, it is important to understand the process and think about what is happening. A computational system may find a function that does not make sense or use a line when the situation suggests that a different model would be more appropriate.

GRADE 8 • UNIT 6 • SECTION B | LESSON 8 355

Lesson Summary

People often collect data in two variables to investigate possible associations between two numerical variables and use the connections that they find to predict more values of the variables. Data analysis usually follows these steps:

1. Collect data
2. Organize and represent the data, then look for an association
3. Identify any outliers and try to explain why these data points are exceptions to the trend that describes the association
4. Find a function that fits the data well

Although computational systems can help with data analysis by graphing the data, finding a function that might fit the data, and using that function to make predictions, it is important to understand the process and think about what is happening. A computational system may find a function that does not make sense or use a line when the situation suggests that a different model would be more appropriate.

Cool-down

Drawing a Line

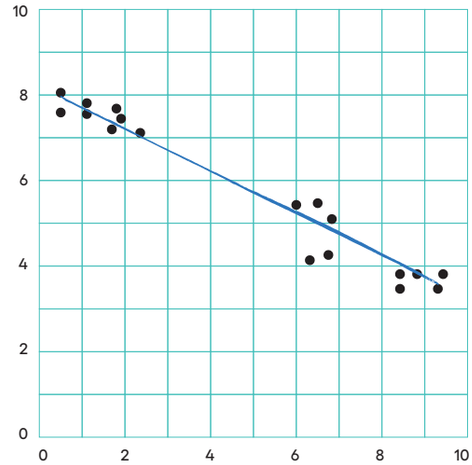
5 min

Students use a given scatter plot to propose a good linear model for the data. They then use that model to predict the value of a new data point within the data.

Student Task Statement

1. Draw a line on the scatter plot that fits the data well.

Sample response:



2. A new point will be added to the scatter plot with $x = 4$. What do you predict for the y-value of this point if it follows the association of the data already in the scatter plot?

Sample response: 6

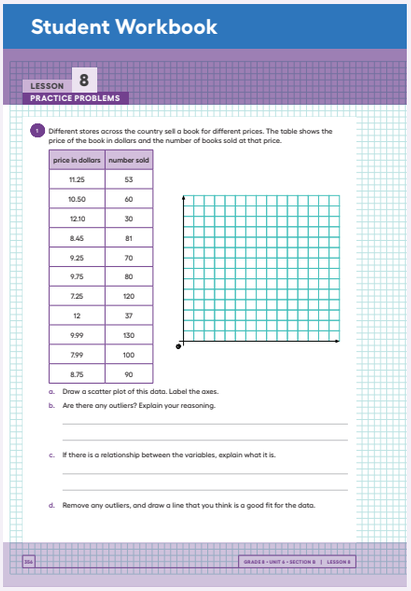
3. A new point will be added to the scatter plot with $x = 10$. What is an example of a y-value of this point if it is considered an outlier?

Sample response: 10

Lesson 8

Practice Problems

3 Problems



Problem 1

Different stores across the country sell a book for different prices. The table shows the price of the book in dollars and the number of books sold at that price.

price in dollars	number sold
11.25	53
10.50	60
12.10	30
8.45	81
9.25	70
9.75	80
7.25	120
12	37
9.99	130
7.99	100
8.75	90

a. Draw a scatter plot of this data. Label the axes.

See image.

b. Are there any outliers? Explain your reasoning.

Yes, at (9.99, 130)

Sample reasoning: This point is much higher than expected on the scatter plot.

c. If there is a relationship between the variables, explain what it is.

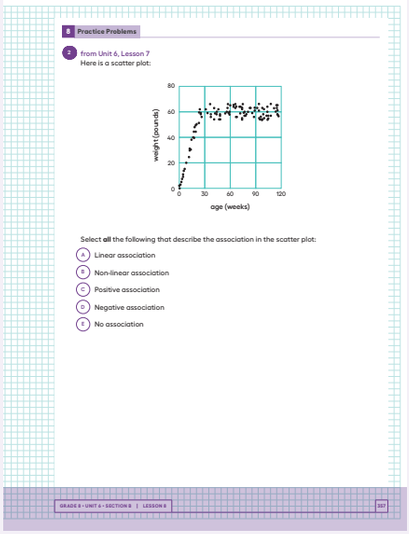
There is a negative linear association between the variables.

Sample reasoning: When the price increases, the number of books sold decreases.

d. Remove any outliers, and draw a line that you think is a good fit for the data.

See image.

Student Workbook


8 Practice Problems

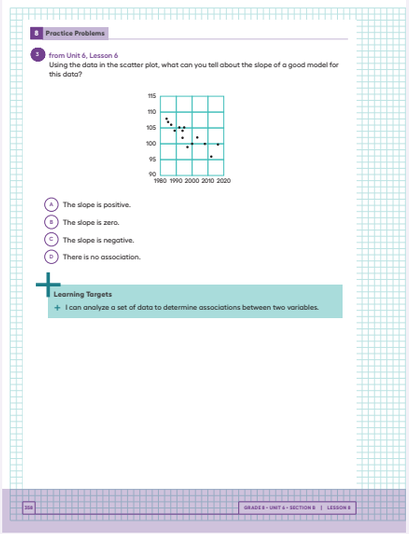
2 from Unit 6, Lesson 7
Here is a scatter plot:

Select **all** the following that describe the association in the scatter plot:

A Linear association
B Non-linear association
C Positive association
D Negative association
E No association

GRADE 8 • UNIT 6 • SECTION B | LESSON 8 227

Student Workbook


8 Practice Problems

3 from Unit 6, Lesson 6
Using the data in the scatter plot, what can you tell about the slope of a good model for this data?

A The slope is positive.
B The slope is zero.
C The slope is negative.
D There is no association.

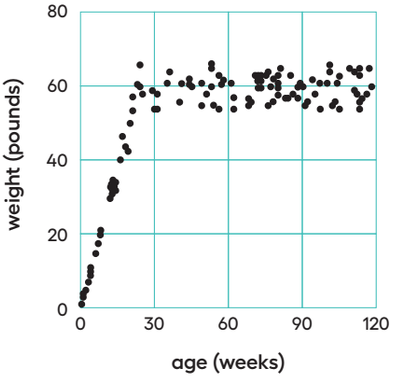
Learning Targets
+ I can analyze a set of data to determine associations between two variables.

228 GRADE 8 • UNIT 6 • SECTION B | LESSON 8

Problem 2

from Unit 6, Lesson 7

Here is a scatter plot:



Select **all** the following that describe the association in the scatter plot:

A. Linear association

B. Non-linear association

C. Positive association

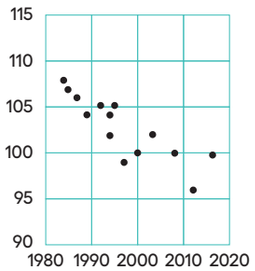
D. Negative association

E. No association

Problem 3

from Unit 6, Lesson 6

Using the data in the scatter plot, what can you tell about the slope of a good model for this data?



A. The slope is positive.

B. The slope is zero.

C. The slope is negative.

D. There is no association.