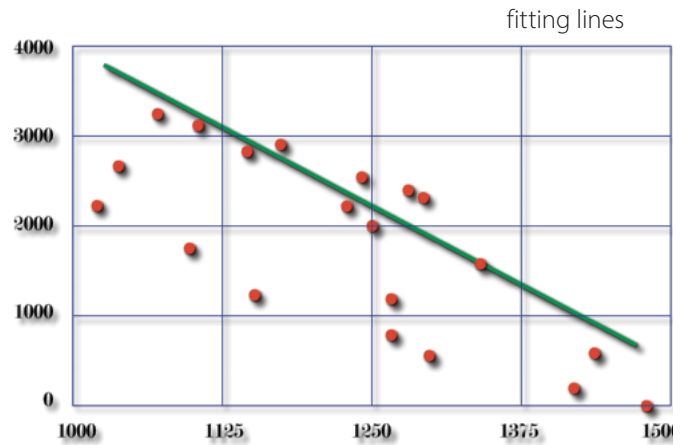


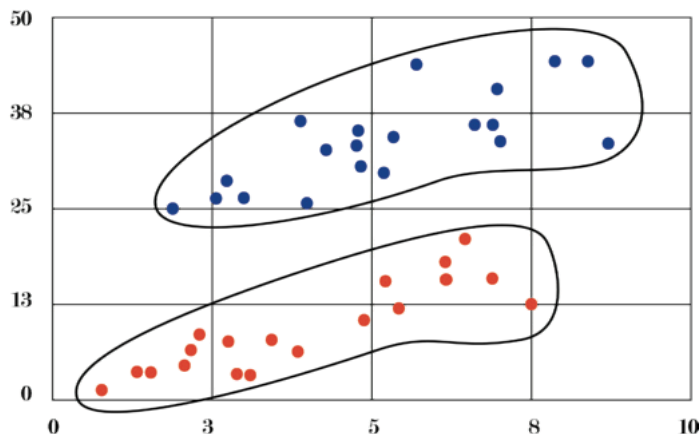
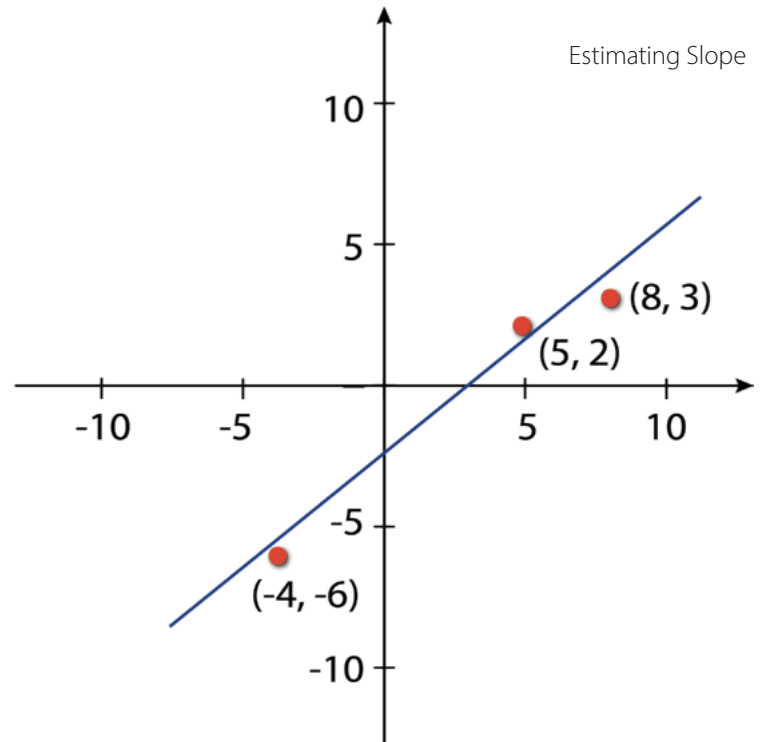
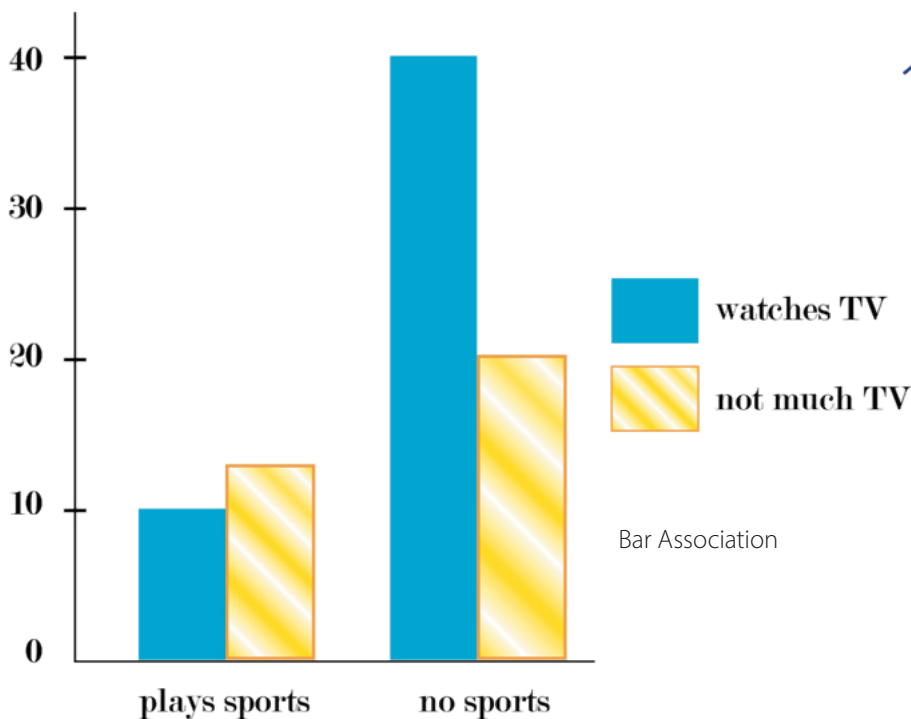
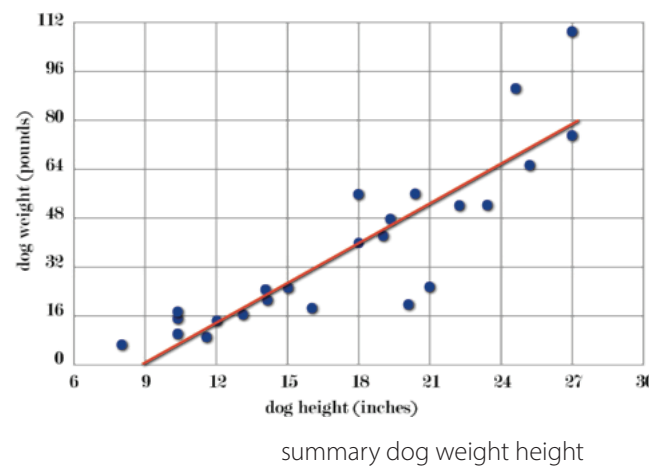


Associations in Data

Teacher Guide



scatter plot



Creative Commons Licensing

This work is licensed under a
Creative Commons Attribution-NonCommercial-ShareAlike
4.0 International License.



You are free:

- to Share**—to copy, distribute, and transmit the work
- to Remix**—to adapt the work

Under the following conditions:

Attribution—You must attribute the work in the following manner:
CKMath 6–8 was originally developed by Open Up Resources and authored by Illustrative Mathematics, <https://www.illustrativemathematics.org>, and is copyrighted as 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). The Open Up Resources 6–8 Math Curriculum is available at: <https://www.openupresources.org/math-curriculum/>.

Adaptations and updates to the IM 6–8 Math English language learner supports and the additional English assessments marked as "B" are copyright 2019 by Open Up Resources and licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Adaptations and updates to the IM K–8 Math Spanish translation of assessments marked as "B" are copyright 2019 by Illustrative Mathematics. These adaptations and updates are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

This particular work is based on additional work of the Core Knowledge® Foundation (www.coreknowledge.org) made available through licensing under a Creative Commons Attribution-Non Commercial-Share Alike 4.0 International License. This does not in any way imply that the Core Knowledge Foundation endorses this work.

Noncommercial—You may not use this work for commercial purposes.

Share Alike—If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

With the understanding that:

For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page:

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Copyright © 2023 Core Knowledge Foundation
www.coreknowledge.org

All Rights Reserved.

Core Knowledge®, Core Knowledge Curriculum Series™, Core Knowledge Math™ and CKMath™ are trademarks of the Core Knowledge Foundation.

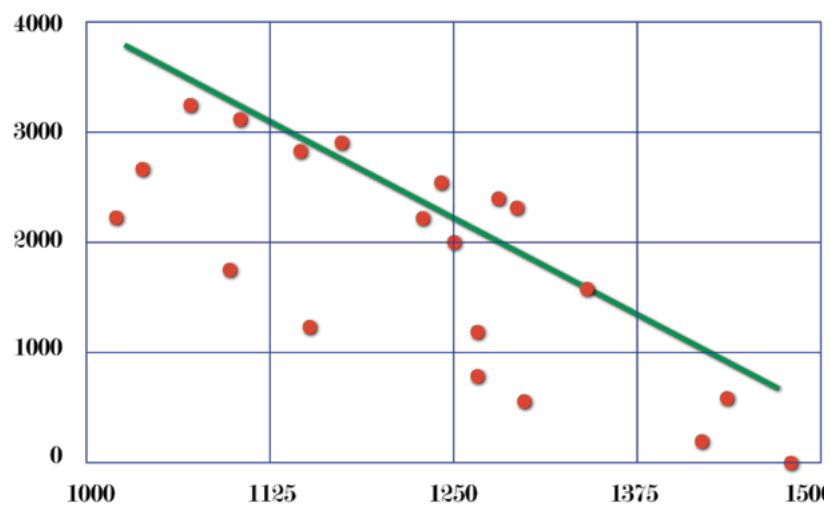
Trademarks and trade names are shown in this book strictly for illustrative and educational purposes and are the property of their respective owners. References herein should not be regarded as affecting the validity of said trademarks and trade names.

ISBN: 979-8-88970-847-6

Associations in Data

Table of Contents

Introduction: Unit Narrative	1
Student Learning Targets	3
Terminology	5
Required Materials	6
Lessons Plans and Student Task Statements:	
Section 1: Lessons 1, 2 Does This Predict That?	7
Section 2: Lessons 3–8 Associations in Numerical Data	40
Section 3: Lessons 9, 10 Associations in Categorical Data	145
Section 4: Lesson 11 Let’s Put It to Work	183
Teacher Resources	191
Family Support Materials	
Unit Assessments	
Assessment Answer Keys	
Cool Downs (Lesson-Level Assessments)	
Instructional Masters	



Associations in Data

Teacher Guide

Core Knowledge Mathematics™

Associations in Data

Unit Narrative

In this unit, students analyze bivariate data—using scatter plots and fitted lines to analyze numerical data, and using two-way tables, bar graphs, and segmented bar graphs to analyze categorical data.

The unit begins with an investigation of a table of data. Measurements of a leg and perimeter of an isosceles right triangle are shown in each row, but column entries are not in order, making it hard to discern a pattern. Students manipulate the data to look for patterns in the table (MP7), then examine a scatter plot of the same data. This motivates the need to use different representations of the same data to find and analyze any patterns.

The second section begins with investigation of two questions: “Are older students always taller?” and “Do taller students tend to have bigger hands?” Students collect data (measurements of each student’s arm span, hand span, and height) and record each student’s measurements together with the student’s age in months. They make a scatter plot for height vs. hand span and select their own methods to display the height data (MP5).

The second section focuses on using scatter plots and fitted lines to analyze numerical data. Students make and examine scatter plots, interpreting points in terms of the quantities represented (MP2) and identifying scatter plots that could represent verbal descriptions of associations between two numerical variables (MP1). They see examples of how a line can be used to model an association between measurements displayed in a scatter plot and they compare values predicted by a linear model with the actual values given in the scatter plot (MP4). They draw lines to fit data displayed in scatter plots and informally assess how well the line fits by judging the closeness of the data points to the lines (MP4). Students compare scatter plots that show different types of associations (MP7) and learn to identify these types, making connections between the overall shape of a cloud of points and trends in the data represented, e.g., a scatter plot of used car price vs. mileage shows a cloud of points that descends from left to right and prices of used cars decrease with increased mileage (MP2). They make connections between the overall shape of a cloud of points, the slope of a fitted line, and trends in the data, e.g., “a line fit to the data has a negative slope and the scatter plot shows a negative association between price of a used car and its mileage.” Outliers are informally identified based on their relative distance from other points in a scatter plot. Students examine scatter plots that show linear and non-linear associations as well as some sets of data that show clustering, describing their differences (MP7). They return to the data on height and arm span gathered at the beginning of the unit, describe the association between the two, and fit a line to the data (MP4).

The third section focuses on using two-way tables to analyze categorical data (MP4). Students use a two-way frequency table to create a relative frequency table to examine the percentages represented in each intersection of categories to look for any associations between the categories. Students also examine and create bar and segmented bar graphs to visualize any associations.

The unit ends with a lesson in which students collect and analyze numerical data using a scatter plot, then categorize the data based on a threshold and analyze the categories based on a two-way table (MP4).

Progression of Disciplinary Language

In this unit, teachers can anticipate students using language for mathematical purposes such as explaining, representing, and interpreting. Throughout the unit, students will benefit from routines designed to grow robust disciplinary language, both for their own sense-making and for building shared understanding with peers. Teachers can formatively assess how students are using language in these ways, particularly when students are using language to:

Explain

- how to estimate using available data (Lesson 1)
- how to use tables and scatter plots to make estimates and predictions (Lesson 3)
- the meaning of slope for a situation (Lesson 6)
- how to use lines to show associations, identify outliers, and answer questions (Lesson 8)

Represent

- data in organized ways (Lesson 1)
- data using two-way tables, bar graphs, and segmented bar graphs (Lessons 9 and 10)
- data using scatter plots (Lesson 11)

Interpret

- situations and graphs involving bivariate data (Lesson 2)
- tables and scatter plots of bivariate data (Lesson 3)
- tables, scatter plots, equations, and situations involving bivariate data (Lesson 4)

In addition, students are expected to compare different representations of the same situation, describe and compare features of scatter plots, justify whether or not lines are good fits for a situation, and justify associations between bivariate data. Students also have opportunities to use language to generalize about what makes a line fit a data set well and generalize about categories for sorting scatter plots.

The table shows lessons where new terminology is first introduced, including when students are expected to understand the word or phrase receptively and when students are expected to produce the word or phrase in their own speaking or writing. Terms from the glossary appear bolded. Teachers should continue to support students' use of a new term in the lessons that follow the one in which it was first introduced.

Learning Targets

Associations in Data

Lesson 1: Organizing Data

- I can organize data to see patterns more clearly.

Lesson 2: Plotting Data

- I can draw a scatter plot to show data that has two paired variables.

Lesson 3: What a Point in a Scatter Plot Means

- I can describe the meaning of a point in a scatter plot in context.

Lesson 4: Fitting a Line to Data

- I can pick out outliers on a scatter plot.
- I can use a model to predict values for data.

Lesson 5: Describing Trends in Scatter Plots

- 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 6: The Slope of a Fitted Line

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

Lesson 7: Observing More Patterns in Scatter Plots

- I can pick out clusters in data from a scatter plot.
- I can use a scatter plot to decide if two variables have a linear association.

Lesson 8: Analyzing Bivariate Data

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

Lesson 9: Looking for Associations

- I can identify the same data represented in a bar graph, a segmented bar graph, and a two-way table.
- I can use a two-way frequency table or relative frequency table to find associations among variables.

Lesson 10: Using Data Displays to Find Associations

- I can create relative frequency tables, bar graphs, and segmented bar graphs from frequency tables to find associations among variables.

Lesson 11: Gone In 30 Seconds

- I can collect data and analyze it for associations using scatter plots, two-way tables, and segmented bar graphs.

lesson	new terminology	
	receptive	productive
8.6.1	scatter plot	
8.6.2	data display attribute	numerical data categorical data
8.6.4	outlier predict overpredict underpredict linear model	
8.6.5	positive association negative association linear association	
8.6.6	nonlinear association no association fitted line	
8.6.7	cluster	
8.6.8		independent variable dependent variable positive association negative association linear association
8.6.9	segmented bar graph relative frequency two-way (frequency) table	
8.6.11		scatter plot outlier cluster

Required Materials

Colored pencils

Copies of Instructional master

Dried linguine pasta

We specified linguine since it is flatter and less likely to roll around than spaghetti.

Measuring tapes

Pre-printed cards, cut from copies of the Instructional master

Rulers marked with centimeters

Stopwatches

Straightedges

A rigid edge that can be used for drawing line segments. Sometimes a ruler is okay to use as a straightedge, but sometimes it is preferable to use an unruled straightedge, like a blank index card.

Section: Does This Predict That?

Lesson 1: Organizing Data

Goals

- Comprehend that a “scatter plot” represents data with two variables and does not represent a function.
- Coordinate (orally and in writing) representations of data in scatter plots and tables.
- Describe (orally and in writing) patterns in representations of data in scatter plots and tables, and use these representations to make predictions.

Learning Targets

- I can organize data to see patterns more clearly.

Lesson Narrative

To open the unit on finding associations between two variables, students are asked to consider a list of unsorted data to notice any patterns or associations. Later, students organize the table to make the pattern more clear and then see a **scatter plot** as a graphic representation of data that can make the association even more obvious. The last activity in the lesson asks students to match different representations of data (MP7) and stresses the importance of labeling each representation so that the meaning is not lost. Scatter plots will be used throughout the unit to help students recognize various associations between variables.

Alignments

Addressing

- 8.SP.A.1: 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.

Building Towards

- 8.SP.A: Investigate patterns of association in bivariate data.

Instructional Routines

- MLR5: Co-Craft Questions
- MLR8: Discussion Supports
- Notice and Wonder

Required Materials

Copies of Instructional master

Required Preparation

Provide one copy of the Instructional master from the activity Tables and Their Scatter Plots for each student.

Student Learning Goals

Let's find ways to show patterns in data

1.1 Notice and Wonder: Messy Data

Warm Up: 5 minutes

The purpose of this warm-up is to elicit useful ideas for the following activities. While students may notice and wonder many things about the table, the most important discussion points are thinking about any associations or patterns.

Building Towards

- 8.SP.A

Instructional Routines

- Notice and Wonder

Launch

Arrange students in groups of 2. Tell students that they will look at a table. Their job is to think of at least one thing they notice and at least one thing they wonder. Display the table for all to see. Ask students to give a signal when they have noticed or wondered about something. Give students 1 minute of quiet think time and then 1 minute to discuss the things they notice with their partner, followed by a whole-class discussion.

Student Task Statement

Here is a table of data. Each row shows two measurements of a triangle.

length of short side (cm)	length of perimeter (cm)
0.25	1
2	7.5
6.5	22
3	9.5
0.5	2
1.25	3.5
3.5	12.5
1.5	5
4	14
1	2.5

What do you notice? What do you wonder?

Student Response

Answers vary. Sample response:

Notice:

- The perimeter increases when the side length increases.
- The perimeters may be rounded to the nearest half unit.
- The data are not in any obvious order.

Wonder:

- Are the measurements exact?
- Is there a pattern in the data?
- Why measure these triangles?

Activity Synthesis

Ask students to share the things they noticed and wondered. Record and display their responses for all to see. If possible, record the relevant reasoning near the table. After each response, ask the class if they agree or disagree and to explain alternative ways of thinking, referring back to the table

each time. If patterns or associations do not come up during the conversation, ask students to discuss these ideas.

1.2 Seeing the Data

10 minutes

In this activity, students continue to look for patterns and associations for variables by sorting the values in a table. They use the table to make predictions about data that is not included in the table (MP2). In the discussion following the activity, students are asked to consider other methods of finding patterns within the data and see a graphical representation that more clearly shows the strong relationship between the variables. The graphical representation can also help improve predictions over those made with only the tables.

Addressing

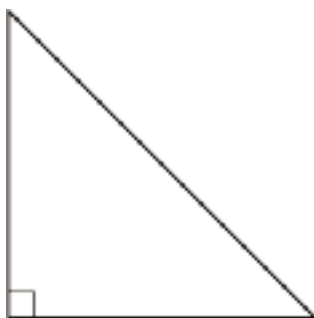
- 8.SP.A.1

Instructional Routines

- MLR8: Discussion Supports

Launch

Keep students in the same groups of 2 from the warm-up. Explain that the data in the warm-up came from fourth grade students who were practicing drawing isosceles right triangles and measuring their perimeters. Display an example of an isosceles right triangle.



Ask students to recall what it means for a triangle to be isosceles and right as well as how to measure the perimeter of a triangle. Ask students if they think there should be a relationship between the length of the two short sides and the entire perimeter. Remind students that specifying 2 sides and the angle between them does determine a unique triangle, so we might expect that knowing the two side lengths and the right angle would be closely related to the length of the perimeter.

Student Task Statement

Here is the table of isosceles right triangle measurements from the warm-up and an empty table.

length of short sides (cm)	length of perimeter (cm)	length of short sides (cm)	length of perimeter (cm)
0.25	1		
2	7.5		
6.5	22		
3	9.5		
0.5	2		
1.25	3.5		
3.5	12.5		
1.5	5		
4	14		
1	2.5		

- How can you organize the measurements from the first table so that any patterns are easier to see? Write the organized measurements in the empty table.
- For each of the following lengths, estimate the perimeter of an isosceles right triangle whose short sides have that length. Explain your reasoning for each triangle.
 - length of short sides is 0.75 cm
 - length of short sides is 5 cm
 - length of short sides is 10 cm

Student Response

Answers vary. Sample responses:

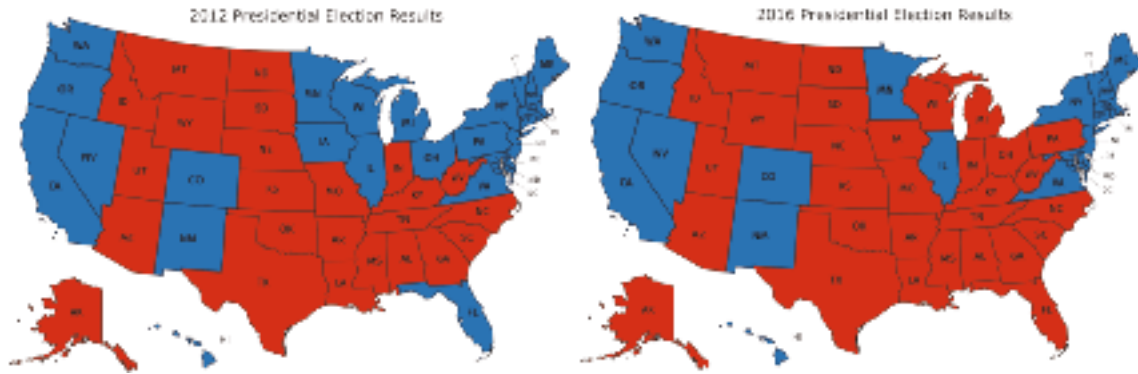
1.

length of short sides (cm)	length of perimeter (cm)
0.25	1
0.5	2
1	2.5
1.25	3.5
1.5	5
2	7.5
3	9.5
3.5	12.5
4	14
6.5	22

2. a. 2.25 cm. It is between the perimeters for the triangles with short lengths 0.5 cm and 1 cm.
- b. 17 cm. As the length of the short sides increases, so does the perimeter. 17 cm is a little less than halfway between 14 cm and 22 cm, the perimeters for triangles with short sides 4 cm and 6.5 cm.
- c. 30 cm. Each time the length of the short side goes up 1 cm, the perimeter goes up about 3 cm, so when the short length is 10 cm, the perimeter should be about 30 cm.

Are You Ready for More?

In addition to the graphic representations of data you have learned, there are others that make sense in other situations. Examine the maps showing the results of the elections for United States president for 2012 and 2016. In red are the states where a majority of electorate votes were cast for the Republican nominee. In blue are the states where a majority of the electorate votes were cast for the Democrat nominee.



1. What information can you see in these maps that would be more difficult to see in a bar graph showing the number of electorate votes for the 2 main candidates?
2. Why are these representations appropriate for the data that is shown?

Student Response

Answers vary. Sample response:

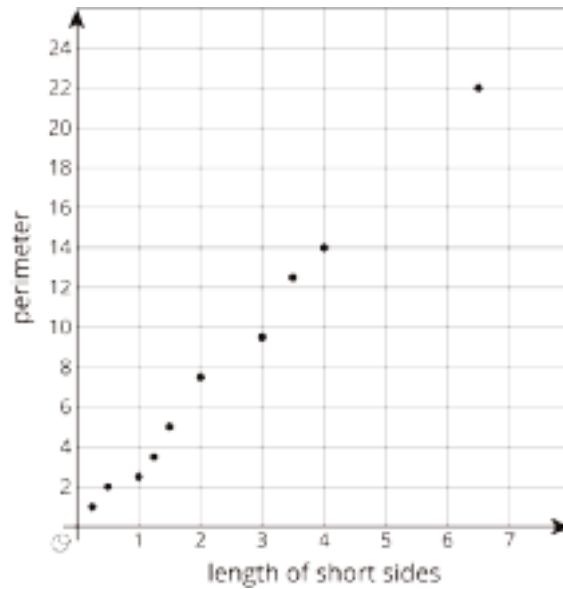
1. The maps show that many of the votes for candidates happen in regions. For example, in the Northeast, they tend to vote for Democrat nominees while in the Southeast, they vote for the Republican nominee.
2. Since electorate votes are cast by state, it makes sense to show the data using a map of the states.

Activity Synthesis

Select students to share their arrangements of the data, the patterns they noticed, and their predictions for the triangles with short side lengths. Display the predictions for all to see.

Ask students to share ideas for other ways to look at the data that might lead to better predictions.

Display a graph of the data for all to see.



To highlight features of the graph, ask:

- "What patterns do you see in the data when it is graphed?" (It is close to linear.)
- "How could you use the graph to estimate the perimeter for the isosceles, right triangles with a short side lengths of 0.75 cm, 5 cm, or 10 cm?" (It is okay for students to not be precise in their answer at this point. Future lessons will give more instruction in how to do this.)

Tell students that this graphical representation of data is called a **scatter plot**. A scatter plot is when two numerical variables are graphed by using one variable as the x -coordinate and the other as the y -coordinate. Data pairs are represented as plotted points.

Note that there is a difference between *time series* graphs and *scatter plots*. In time series graphs, a single variable is recorded at multiple time points and plotted on a graph. In a scatter plot, two variables are measured and plotted on a graph. For scatter plots, time may be one of the variables, but it should be possible to have more than one measurement for the second variable for the same time measurement. For example, when comparing the price of a car to its model year uses the year as one of the variables, but two cars made in the same year could have two different prices. The price for a single car throughout different years would be represented in a time series graph since it only has one price at any given time.

Access for English Language Learners

Speaking: MLR8 Discussion Supports. As students describe the patterns they noticed and their predictions, revoice student ideas to demonstrate mathematical language use by restating a statement as a question (e.g., “As you estimated the perimeters, how did you use the length of the triangles’ short side to help you estimate?”) in order to clarify, apply appropriate language, and involve more students.

Design Principle(s): Support sense-making; Optimize output (for explanation)

1.3 Tables and Their Scatter Plots

15 minutes

An essential part of creating and understanding scatter plots is interpreting the meaning of the points (MP2). In this activity, students match tables of data with scatter plots representing the same information (MP7). After matching appropriately, students are asked to include titles for the axes of the scatter plots. Following the activity, the importance of the axis labels is discussed.

Addressing

- 8.SP.A.1

Instructional Routines

- MLR5: Co-Craft Questions

Launch

Arrange students in groups of 2. Distribute 1 copy of the tables from the Instructional master to each group and resolve any clarifying questions about the data in the tables. In particular, students may wish to know about particular terminology. For example, “kilowatt hours” are a unit of electrical energy that most electricity companies use to measure how much electricity a customer used to determine how much to charge them. Another example is “carat” which is a unit of weight equal to 200 milligrams used for measuring precious stones.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. To support development of organizational skills, check in with students within the first 2–3 minutes of work time. Look for students who have matched at least one table with one of the scatter plots accurately. Check to make sure students are in agreement with their matches and either person can justify the decision.

Supports accessibility for: Memory; Organization

Access for English Language Learners

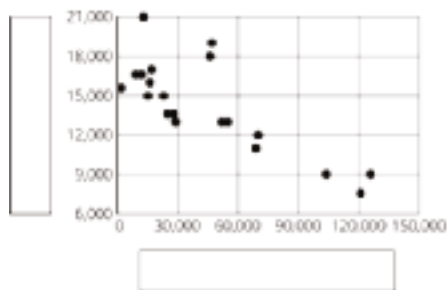
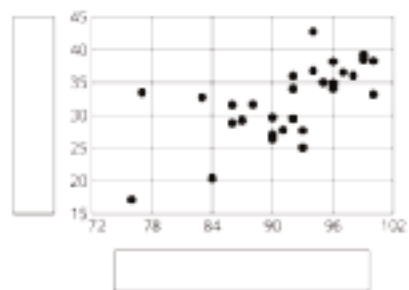
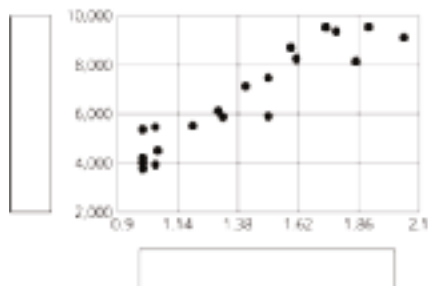
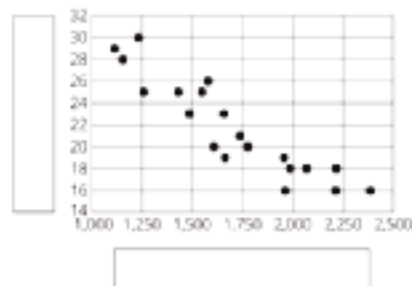
Conversing, Writing: MLR5 Co-Craft Questions. Display only the four graphs without revealing the questions that follow. Invite pairs of students to write mathematical questions about the graph. Then, invite 2–3 groups to share their questions with the class. Look for questions that ask students to make sense of the relationship between the two quantities represented on each axis. Next, reveal the questions of the activity. This routine allows students to produce the language of mathematical questions and talk about the relationship between quantities that are represented graphically.

Design Principle(s): Maximize meta-awareness; Support sense-making

Student Task Statement

Here are four **scatter plots**. Your teacher will give you four tables of data.

- Match each table with one of the scatter plots.
- Use information from the tables to label the axes for each scatter plot.



Student Response

1. (top left) Title: Used car price vs. mileage, horizontal axis: mileage (miles), vertical axis: price (dollars)
2. (top right) Title: August daily high vs. energy consumed, horizontal axis: daily high (degrees Fahrenheit), vertical axis: energy consumed (kilowatt hours)

3. (bottom left) Title: Diamond prices vs. size in carats, horizontal axis: weight (carats), vertical axis: price (dollars)

4. (bottom right) Title: New car weight vs. fuel efficiency, horizontal axis: mass (kilograms), vertical axis: fuel efficiency (miles per gallon)

Activity Synthesis

The goal of this discussion is for students to develop strategies for relating different representations of data as well as to see the importance of labeling and numbering the axes.

To highlight the axis numbering in preparation for later lessons where students will draw their own scatter plots and need to select a good interval and range, ask:

- "What are some things you notice about the numbers on the axes for the different scatter plots?" (They don't all start at 0 and go up by 1s, 5s, or 10s.)
- "Why do you think it makes sense for the graph comparing temperature and energy to start at 72 and go to 102?" (The minimum temperature is 77 and the maximum temperature is 100.)
- "Why do the numbers for the graph comparing temperature and energy go up by 6 rather than 1 or 10." (The numbers would be too close together if it were 1 and too far apart if they were 10.)

Ask partners to select one of the other axes in the graphs (including vertical axes) and discuss why they may have been chosen the way they were. Select 1–2 groups to share their ideas with the class.

Tell students that they should think about the maximum and minimum values as well as the range (the distance between the maximum and minimum values) when setting the scale for the different axes. For the graphs in this unit, it is not usually essential to include the point (0, 0) in the graph, so that makes the axis labels even more important.

To conclude the discussion, consider asking some of the following questions:

- "How did you match the tables to the scatter plots?"
- "Why is it important to include labels for the axes on scatter plots?" (So that any patterns that are found can be recognized from the scatter plot without having to go back to the table.)
- "Why is it important to include units in the axis labels?" (So that the patterns found can be understood easily)
- "The same data is presented in the tables as in the scatter plots. Which is easier to understand? Explain your reasoning."

Lesson Synthesis

To highlight the progression of representations seen today (unorganized table, ordered table, scatter plot) that help to highlight any patterns that may be present in data, ask:

- "Why would sorting the information in the table be helpful?" (It makes it easier to see the minimum and maximum values of the value you sort by.)
- "When looking for relationships between two variables, what are some graphic representations that might be helpful to use?" (An organized table helps show trends of one variable while another goes from smallest to largest. A scatter plot also shows how one variable changes in relation to the other.)
- "What is a scatter plot and how is it different from plotting points for a function?" (For a function, every input has a single output, but a scatter plot uses data that does not have an input or output and a single value for one variable could have multiple values in the second variable.)

Remind students that, when reorganizing data, it is important to continue to label what the information represents. A table without titles or a scatter plot without labels may show some relationship between numbers, but is meaningless outside of the context.

1.4 Squashed Spheres

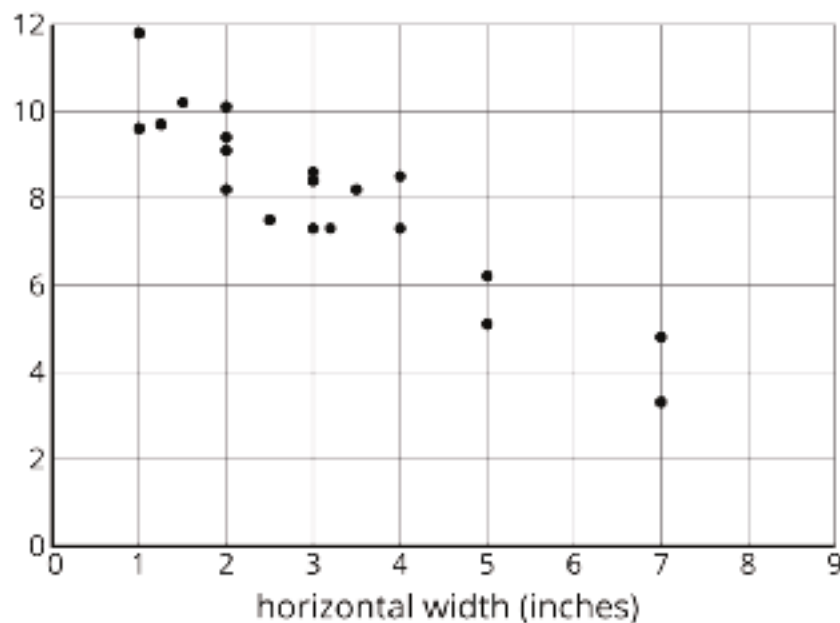
Cool Down: 5 minutes

Addressing

- 8.SP.A.1

Student Task Statement

Twenty rubber spheres were compressed with varying amounts of force. The widths and heights of the resulting shapes were measured. Here is a scatter plot that shows the measurements for each sphere.



1. Label the vertical axis of the scatter plot.

2. If a compressed rubber sphere has a 6-inch width, is its height closer to 5 inches or to 11 inches? Explain your reasoning.

Student Response

1. Vertical height (inches)
2. 5 inches. When the width increases, the height tends to decrease. To keep in line with the rest of the data, the height should be closer to 5 inches than 11 inches.

Student Lesson Summary

Consider the data collected from pulling back a toy car and then letting it go forward. In the first table, the data may not seem to have an obvious pattern. The second table has the same data and shows that both values are increasing together.

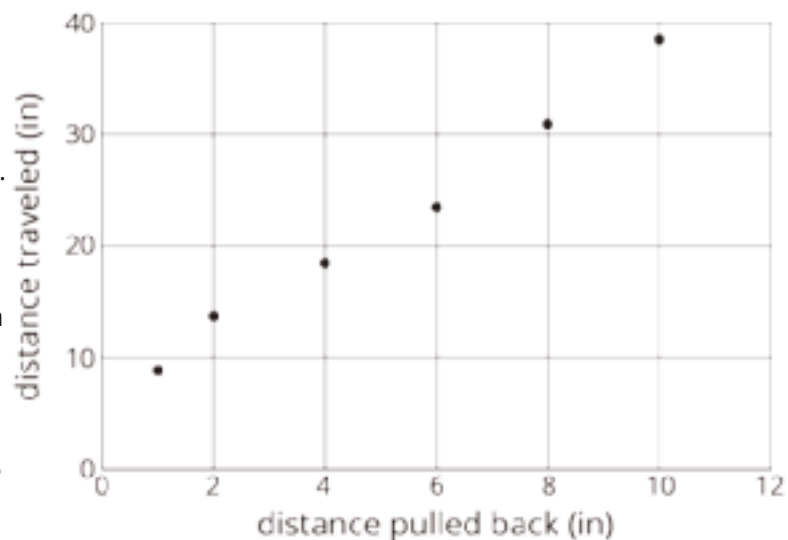
Unorganized table

distance pulled back (in)	distance traveled (in)	distance pulled back (in)	distance traveled (in)
6	23.57	1	8.95
4	18.48	2	13.86
10	38.66	4	18.48
8	31.12	6	23.57
2	13.86	8	31.12
1	8.95	10	38.66

Organized table

A scatter plot of the data makes the pattern clear enough that we can estimate how far the car will travel when it is pulled back 5 inches.

Patterns in data can sometimes become more obvious when reorganized in a table or when represented in **scatter plots** or other diagrams. If a pattern is observed, it can sometimes be used to make predictions.





Glossary

- scatter plot

Lesson 1 Practice Problems

Problem 1

Statement

Here is data on the number of cases of whooping cough from 1939 to 1955.

year	number of cases
1941	222,202
1950	120,718
1945	133,792
1942	191,383
1953	37,129
1939	103,188
1951	68,687
1948	74,715
1955	62,786
1952	45,030
1940	183,866
1954	60,866
1944	109,873
1946	109,860
1943	191,890
1949	69,479
1947	156,517

- Make a new table that orders the data by year.
- Circle the years in your table that had fewer than 100,000 cases of whooping cough.
- Based on this data, would you expect 1956 to have closer to 50,000 cases or closer to 100,000 cases?

Solution

a.

year	number of cases
1939	103,188
1940	183,866
1941	222,202
1942	191,383
1943	191,890
1944	109,873
1945	133,792
1946	109,860
1947	156,517
1948	74,715
1949	69,479
1950	120,718
1951	68,687
1952	45,030
1953	37,129
1954	60,886
1955	62,786

- b. The years 1948, 1949, 1951, 1952, 1953, 1954, and 1955 had fewer than 100,000 cases of whooping cough.
- c. This data seems to show the number of cases decreasing over time, so I would expect 1956 to have closer to 50,000 cases than 100,000.

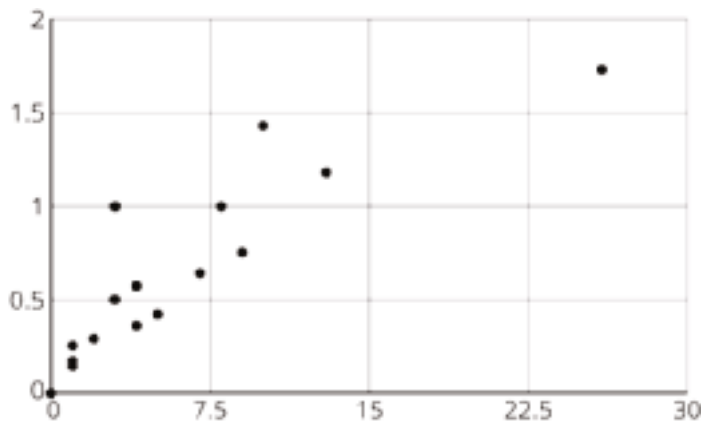
Problem 2

Statement

In volleyball statistics, a block is recorded when a player deflects the ball hit from the opposing team. Additionally, scorekeepers often keep track of the average number of blocks

a player records in a game. Here is part of a table that records the number of blocks and blocks per game for each player in a women's volleyball tournament. A scatter plot that goes with the table follows.

blocks	blocks per game
13	1.18
1	0.17
5	0.42
0	0
0	0
7	0.64



Label the axes of the scatter plot with the necessary information.

Solution

The horizontal axis should be labeled “blocks,” and the vertical axis should be labeled “blocks per game.”

Problem 3

Statement

A cylinder has a radius of 4 cm and a height of 5 cm.

- What is the volume of the cylinder?
- What is the volume of the cylinder when its radius is tripled?
- What is the volume of the cylinder when its radius is halved?

Solution

- $80\pi \text{ cm}^3$
- $720\pi \text{ cm}^3$
- $20\pi \text{ cm}^3$

(From Unit 5, Lesson 18.)

Lesson 2: Plotting Data

Goals

- Create a representation of single-variable data using a box plot, histogram, or dot plot, and compare and contrast (orally) these representations with a scatter plot.
- Create a scatter plot from a table of data, and describe (orally and in writing) the trend of the data.
- Create a table of collected data, and explain (orally) how to organize the data.

Learning Targets

- I can draw a scatter plot to show data that has two paired variables.

Lesson Narrative

In prior grades, students represented the distribution of a single statistical variable using dot plots, histograms, and box plots. In this lesson, students review those ways of displaying data and compare them with representing the relationship between *two* variables in a scatter plot. They aggregate data about hand span, arm span, and height for all students in the class, and use the data to create plots of a single variable as well as a scatter plot. (The data will be used again in later lessons, so should be kept in a spreadsheet or a permanent visual display in the classroom.) They notice that scatter plots can convey information about the relationship between two variables that representations of each variable separately do not reveal. They understand that every point in a scatter plot represents two measures of a single individual in the population.

In this lesson, students choose an appropriate way to display a data set (MP5) and see that the structure in a scatter plot can reveal information about a data set that is not visible in other representations (MP7).

Alignments

Building On

- 6.SP.B.4: Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

Addressing

- 8.SP.A.1: 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.

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect
- MLR7: Compare and Connect

- MLR8: Discussion Supports

Required Materials

Measuring tapes

Rulers marked with centimeters

Required Preparation

For the Gathering Data activity, each pair of students will need to measure some lengths in centimeters. Rulers, meter sticks, or tape measures would work for this purpose, so long as they are marked in centimeters.

At the end of the Gathering Data activity, students are instructed to add their data to a class table. It would be fine for students to record their individual data on the board for all to see. However, this data will be used again in a later lesson, so you need a way of preserving it for later. A useful mechanism might be to set up a shared spreadsheet that students access through a browser.

Student Learning Goals

Let's collect and display some data about the class.

2.1 Representing Data

Warm Up: 5 minutes

In prior grades, students studied different ways of representing sets of data. The purpose of this warm-up is for students to recall ways that two different sets of data can be represented. Data visualization is very useful for understanding patterns that are not visible in other ways. Different representations can highlight different aspects of the data and lead the viewer to see different patterns (MP7).

Students are given two scenarios and asked for appropriate representations. One scenario contains quantitative data (distance run) while the other has categorical data (favorite color). Each scenario can be represented in more than one way. As students work, monitor for different students using a variety of display choices.

Building On

- 6.SP.B.4

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect

Launch

Give students 2 minutes of quiet work time followed by a whole-class discussion.

Ask students to brainstorm different graphical representations of data they have used or seen in the past. Record and display answers for all to see.

Student Task Statement

Lin surveyed 30 students about the longest time they had ever run. Andre asked them about their favorite color. How could Lin and Andre represent their data sets? Would they represent them in the same way? Why or why not?

Student Response

Answers vary. Sample response: For the longest run, Lin could use a dot plot, a histogram, or a box plot. For the colors, Andre can use a bar graph or a circle graph. They represent the data in different ways because Lin's data is quantitative and Andre's is categorical.

Activity Synthesis

The purpose of the discussion is for students to think about different representations of data and to emphasize the difference in categorical and numerical data types.

Select previously identified students to share their responses and record for all to see. To highlight the differences between the representations, after each student shares ask:

- "What patterns would Lin be able to see if she represented her data in that way?" (A dot plot would show all of the data, so would be good for showing details. A histogram would give a good overall picture of the data, but would hide details. A box plot would also be useful for showing an overview of the data and would highlight the median, but would not show the individual data values.)
- "What patterns would Andre be able to see if he represented his data in that way?" (A bar graph would be useful for quickly seeing the relative order of color preferences. A circle graph would be useful for quickly seeing how preferred each color is when compared to the whole.)

After the responses have been shared, ask, "Why couldn't Lin and Andre use the same graph type to represent their data?" (Lin has numerical data and Andre has categorical data.)

2.2 Gathering Data

10 minutes

Data collection is an important part of any statistics unit. Here, students collect data about a partner using the appropriate tools (MP5) and submit the data for future use.

In a lesson later in the unit, students analyze the class data they collect here using the methods they learn in the next few lessons. A copy of the class's data should be kept. It may be convenient to set up a shared spreadsheet if students have access to Internet-enabled devices.

Note: Some students may be self-conscious about their body measurements. On the other hand, it can be motivating for students to use their own measurements to analyze statistically. Depending on your class, consider collecting the data and analyzing a similar data set in a future lesson (measurements from the staff, a different class, or invented data similar to the data collected).

Addressing

- 8.SP.A.1

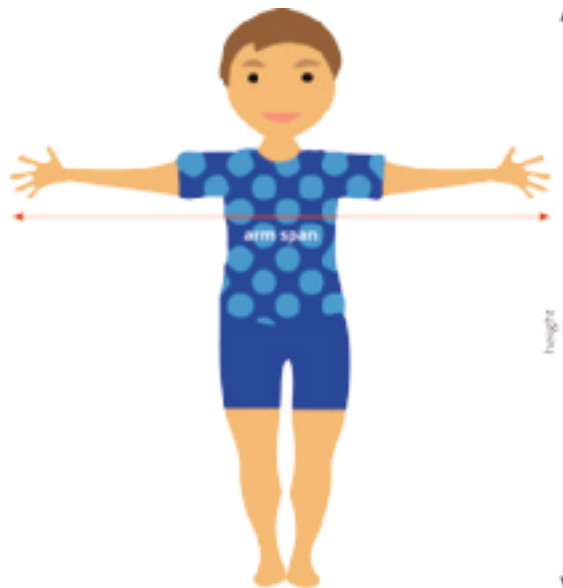
Instructional Routines

- MLR8: Discussion Supports

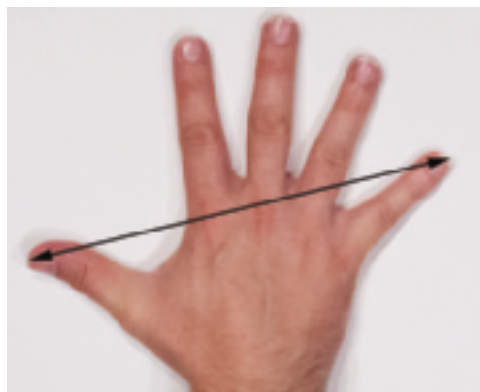
Launch

Arrange students in groups of 2. Provide access to rulers marked in centimeters.

It may be helpful to demonstrate how to measure arm span as defined in the task. An image is included here to aid understanding.



It may also be useful to demonstrate how to measure hand span as defined in the task. An image of this is also included.



Prepare a table with enough rows for the entire class to include their data for all to see. The data will be needed again in a future lesson, so it might be helpful to keep a copy digitally, on a large piece of chart paper, or something that could be brought out again later. It is important to collect

this data so that the class will have enough data to look for patterns later, but the data does not need to be connected to individuals, so collecting names attached to the data is not important.

As groups complete their tables, select one student from each group to include their group's data in the table for the entire class. Each student should also have access to the class data for the next activity as well as a for a future lesson.



Access for Students with Disabilities

Engagement: Develop Effort and Persistence. Provide prompts, reminders, guides, rubrics, or checklists that focus on increasing the length of on-task orientation in the face of distractions. For example, create a checklist for students to complete as they measure each person’s height, arm span, and hand span and record the data in the table.

Supports accessibility for: Attention; Social-emotional skills



Student Task Statement

Are older students always taller? Do taller students tend to have bigger hands? To investigate these questions, the class will gather data.

- A person’s *arm span* is the distance between the tips of their index fingers, when their arms are fully spread out.
- A person’s *hand span* is the distance from the tip of their thumb to the tip of their little finger, when their fingers are fully spread out.

1. Each partner should:
- Measure the other partner’s height, arm span, and hand span for their right hand to the nearest centimeter.
 - Record the other partner’s measurements and age (in months) in the table.

	height (cm)	arm span (cm)	hand span (cm)	age (months)
partner A				
partner B				

2. One partner records the data from your table in a table of data for the entire class.

Student Response

Answers vary. Sample response:

	height (cm)	arm span (cm)	hand span (cm)	age (months)
partner A	160	157	19	162, because 13 years is 156 months, plus 6 more months.
partner B	152	155	18	165

Activity Synthesis

The goal for this activity is for students to see the importance of careful data collection as well as an organizational strategy for keeping up with the data as it is collected.

Consider asking some of the following questions:

- "Do you notice any patterns in the data from the table?"
- "Is the class data organized in a way that is useful for noticing patterns?" (Not likely.)
- "How could we reorganize the table to make it more useful?" (Sort from shortest to tallest or youngest to oldest.)
- "What questions do you have about the data?" (Are some of the items related to one another? Do height and hand span have any relationship?)

Access for English Language Learners

Speaking: MLR8 Discussion Supports. As students describe what they notice in the table, revoice student ideas to demonstrate mathematical language use. Press for details in students' explanations by requesting that students challenge an idea, elaborate on an idea, or give an example. This will help students to produce and make sense of the language needed to communicate their own ideas, and identify the importance of careful data collection.

Design Principle(s): Support sense-making; Optimize output (for explanation)

2.3 Scatter Plots

20 minutes (there is a digital version of this activity)

In this activity, students choose appropriate ways to display data using data that comes from the items students measured in the previous activity (MP4).

After drawing the distributions of the individual variables and then the scatter plot, students look for patterns in the data that might not have been visible in the tabular format.

Addressing

- 8.SP.A.1

Instructional Routines

- MLR7: Compare and Connect

Launch

Keep students in groups of 2. For the sake of time, you may tell students that groups should discuss a plan for working on problems 1 and 2, then each student can create the display individually.

Brainstorm the ways that students can display the distribution of height data. Possibilities include dot plots, histograms, and box plots. Remind students to label axes and include units of measurement (MP6).

If using the digital materials: The digital tools allow students to input the data and quickly view the difference between many graphical representations. Have students discuss the benefits of different graphs for this specific data (height).

Access for Students with Disabilities

Action and Expression: Provide Access for Physical Action. Provide access to tools and assistive technologies such as a the digital materials or graphing software. Some students may benefit from a checklist or list of steps to be able to use the digital materials or software.

Supports accessibility for: Organization; Conceptual processing; Attention

Anticipated Misconceptions

Some students may struggle to find a good way to number the axes so that the data is visible, but not misleading. As students have already seen in scatter plots for this unit, it is not essential to start from 0 on a scatter plot. To accurately show the the data, ask students to find the minimum and maximum values in the data set and use those to help think about reasonable boundaries for the left, right, bottom, and top sides of the scatter plot. The increments should be chosen based on the minimum and maximum values of the boundaries.

For example, if the minimum value is 8,532 and the maximum value is 13,456, then the left side of the graph might be 8,000 or 8,500 and the right side 13,500 or 14,000. The grid divisions should not represent steps of 1 (There would be thousands of grid lines!), but rather 300 or so. These values allow all of the data points to fit without them being too clustered together and indistinguishable.

In most cases, estimates of position can be used to plot the points in a scatter plot, but using technology makes it more precise.

Student Task Statement

1. What types of graphical representations could be used to show the class's height measurements? Make a graphical representation of the class's height measurements.
2. Choose a color and use it to plot a point on the coordinate plane that represents your own height and hand span. Then, in the same color, plot a second point that represents your partner's height and hand span.



3. In a different color, plot the height and hand span of each student in your class, making a scatter plot of the heights and hand spans for the entire class.
4. Based on your scatter plot, answer these questions:
 - a. Do taller students in your class tend to have bigger hands? Explain how you know.
 - b. Is hand span a linear function of height? Explain how you know.

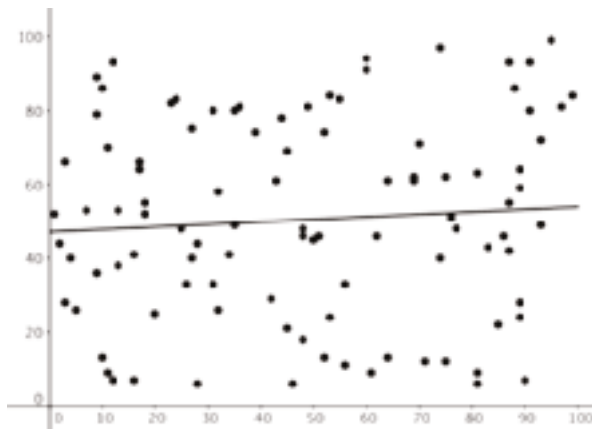
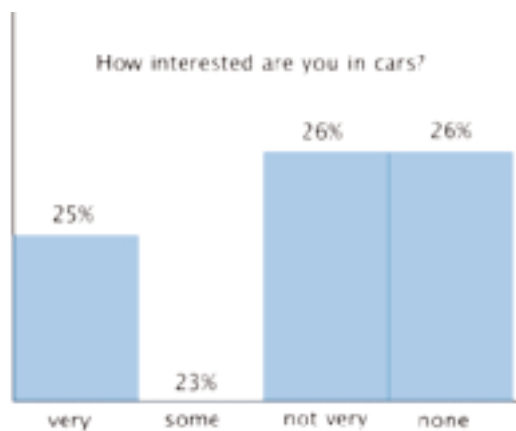
Student Response

1. Answers vary. Sample response. Histogram, a dot plot, or a box graph.
2. Graphs vary based on class data. The graph should be labeled with height, in cm, along one axis and hand spans, in cm, along the other axis. If the heights are along the horizontal axis, then each student should be represented by a single point with coordinates in the form of (height, hand span).
3. Graphs vary based on class data.

4.
 - a. Yes, taller students tend to have greater hand spans. On the graph, as the heights increase, the hand spans tend to increase as well.
 - b. No, the hand span is not a linear function of height, because I cannot draw a straight line through all of the data points. Sometimes there are different hand span measurements for students with the same height.

Are You Ready for More?

Although the data may be accurate, displaying the data incorrectly can tell the wrong story. What is wrong with each of these graphic representations of the data?



Student Response

Answers vary. Sample response: In the scatter plot, the axes are not labeled, so it is not clear what the data represents. The line also doesn't seem to make sense with the rest of the points. In the bar graph, the heights of the bars do not match the values written above them. All four responses are pretty close to one another, but the bars make it look like there is a major difference.

Activity Synthesis

The goal of the discussion is for students to explore appropriate data representations including one of the main focuses of the unit: scatter plots.

Select students to share their graphs for height measurements and their scatter plots for heights and hand spans in the class. In particular, look for students who chose to represent the data in a different way.

For example, to graph height data, students could group different heights into bins to graph with a histogram or graph each student's height individually on a dot plot.

Additionally, for the scatter plot, height could be along the x -axis and hand span along the y -axis or this could be reversed. If no students have tried these methods, mention them as a possibility. Use MLR 7 (Compare and Connect) to ask what is similar and what is different about the graph when height and hand span are on different axes.

Select a single student's information from the data and ask students to identify where that person's data is represented in each of the displays.

Make sure that students understand the distinction between representations of the distribution of a *single* variable (height data alone) and this new representation that contains information about two variables at once (height and hand span). Tell students that we will be using scatter plots to look at the relationship between different variables that cannot be seen as easily when we represent the distributions separately.

Access for English Language Learners

Speaking, Listening: MLR7 Compare and Connect. As students share their visual displays, ask students to identify what is the same and what is different about each representation. Draw students' attention to the ways individual data points were represented in each of the displays. In this discussion, emphasize the mathematical language used to make sense of the different graphical representations to display the data. These exchanges strengthen students' mathematical language use and reasoning used to represent data.

Design Principle(s): Maximize meta-awareness

Lesson Synthesis

Compare and contrast the different kinds of representations of data that students worked with in this lesson. Some represent a single categorical variable. Some a single quantitative variable. Scatter plots show the relationship between *two* variables. Note that we will be using scatter plots extensively in this unit to investigate the relationship between different variables in many different contexts.

To highlight the different representations students created and saw today, ask:

- "When do we use histograms or box plots to represent a data set? When do we use a scatter plot?" (Histograms or box plots are useful when displaying numerical data from a single variable. Scatter plots are more useful when looking for a relationship between two variables that produce numerical data.)
- "What does a point in a scatter plot represent?" (Two measures about a single person or object.)

2.4 Right Side Measurements

Cool Down: 5 minutes

Students create a scatter plot from data given in a given table. Then they circle the data point in the scatter plot that corresponds to a particular entry in the table, reinforcing the connection between these representations.

Addressing

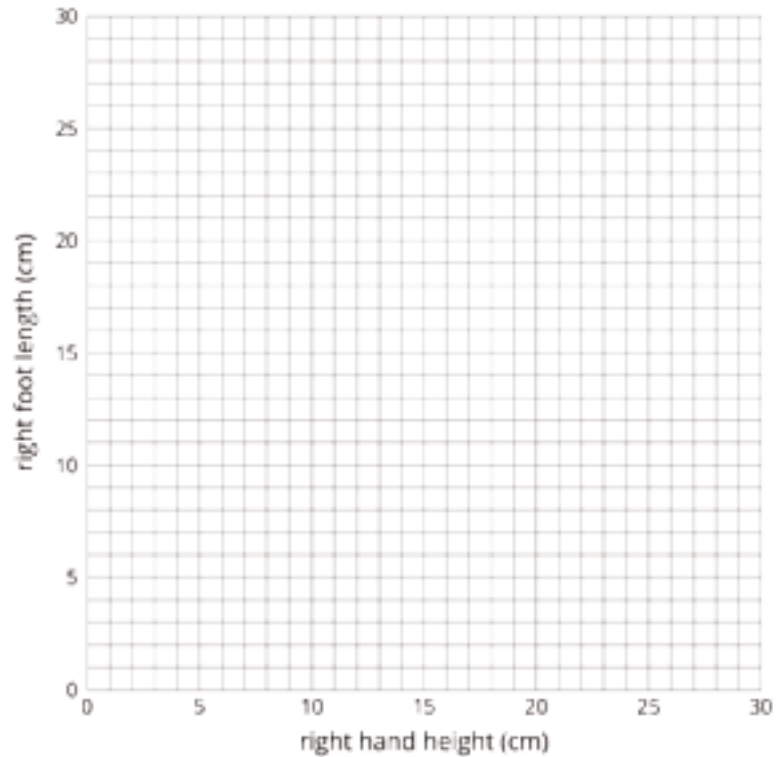
- 8.SP.A.1

Student Task Statement

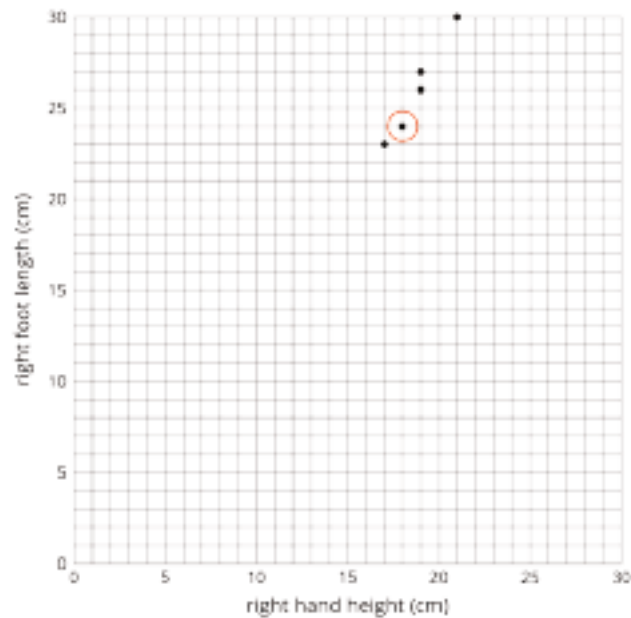
The table shows measurements of right hand length and right foot length for five people.

	right hand length (cm)	right foot length (cm)
person A	19	27
person B	21	30
person C	17	23
person D	18	24
person E	19	26

1. Draw a scatter plot for the data.
2. Circle the point in the scatter plot that represents Person D's measurements.



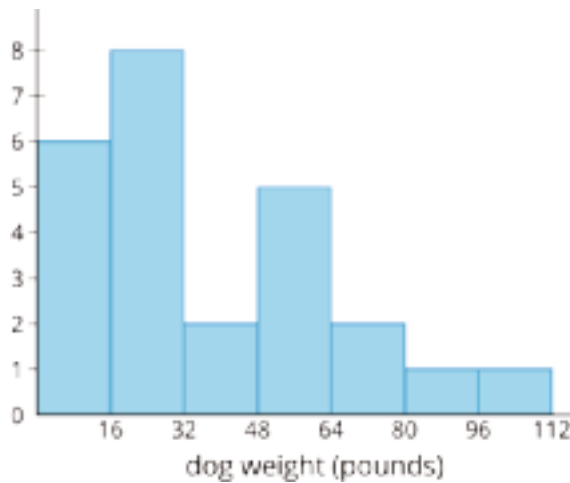
Student Response



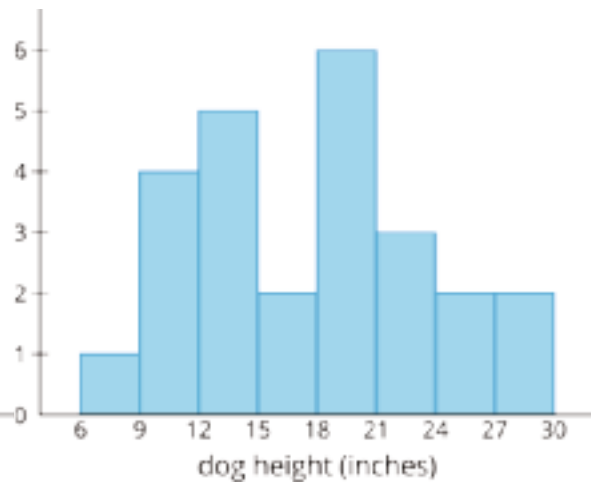
Student Lesson Summary

Histograms show us how measurements of a single attribute are distributed. For example, a veterinarian saw 25 dogs in her clinic one week. She measured the height and weight of each dog.

This histogram shows how the weights of the dogs are distributed.

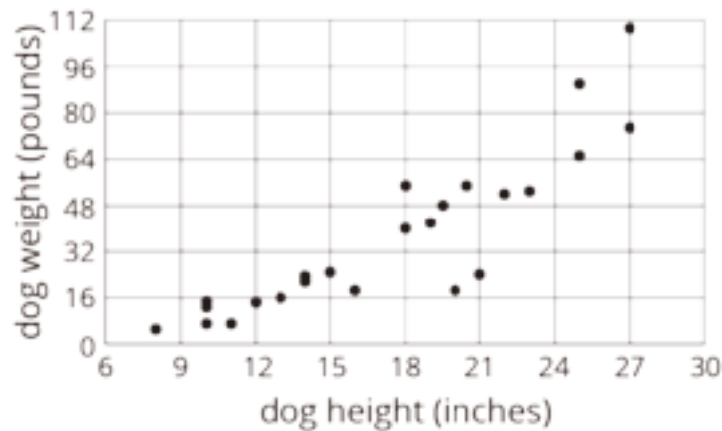


This histogram shows how the heights of the dogs are distributed.



These histograms tell us how the weights of the dogs and how the heights of dogs were distributed. But, they do not give any evidence of a connection between a dog's height and its weight.

Scatter plots allow us to investigate possible connections between two attributes. In this example, each plotted point corresponds to one of the 25 dogs, and its coordinates tell us the height and weight of that dog. Examination of the scatter plot allows us to see a connection between height and weight for the dogs.



Lesson 2 Practice Problems

Problem 1

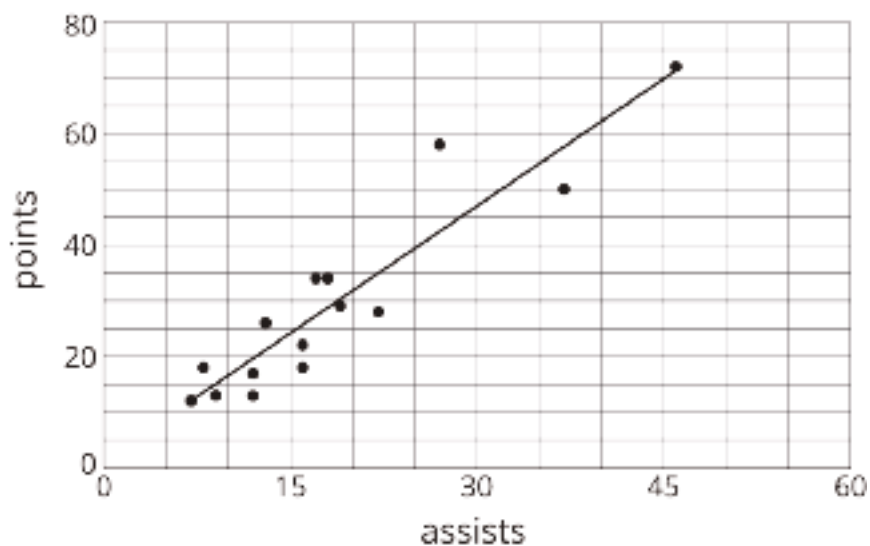
Statement

In hockey, a player gets credited with a “point” in their statistics when they get an assist or goal. The table shows the number of assists and number of points for 15 hockey players after a season.

assists	points
22	28
16	18
46	72
19	29
13	26
9	13
16	22
8	18
12	13
12	17
37	50
7	12
17	34
27	58
18	34

Make a scatter plot of this data. Make sure to scale and label the axes.

Solution



Problem 2

Statement

Select **all** the representations that are appropriate for comparing bite strength to weight for different carnivores.

- A. Histogram
- B. Scatter plot
- C. Dot plot
- D. Table
- E. Box plot

Solution

["B", "D"]

Problem 3

Statement

When is it better to use a table? When is it better to use a scatter plot?

Solution

Answers vary. Sample response: Scatter plots are best when looking for an overall pattern (or lack of one). Tables are best when looking for the precise details of the data.

Problem 4

Statement

There are many cylinders with radius 6 meters. Let h represent the height in meters and V represent the volume in cubic meters.

- Write an equation that represents the volume V as a function of the height h .
- Sketch the graph of the function, using 3.14 as an approximation for π .
- If you double the height of a cylinder, what happens to the volume? Explain this using the equation.
- If you multiply the height of a cylinder by $\frac{1}{3}$, what happens to the volume? Explain this using the graph.

Solution

- $V = 36\pi h$
- The graph is a line starting from $(0, 0)$ then through about $(1, 113)$ and $(2, 226)$.
- If you double the height, the volume doubles. Replacing h with $2h$ in the equation gives $V = 36\pi \cdot 2h = 2(36\pi h)$, double the original volume.
- If you multiply the height by $\frac{1}{3}$, the volume is also multiplied by $\frac{1}{3}$. On the graph this can be seen using similar triangles, or by noting the relationship is proportional.

(From Unit 5, Lesson 17.)

Section: Associations in Numerical Data

Lesson 3: What a Point in a Scatter Plot Means

Goals

- Coordinate (orally and in writing) data in a table and points on a scatter plot.
- Describe (orally) the trend of the data, and use the trend to predict unknown values.
- Interpret (orally and in writing) a point on a scatter plot in context.

Learning Targets

- I can describe the meaning of a point in a scatter plot in context.

Lesson Narrative

In this lesson, students continue their investigation of scatter plots. They interpret points in a scatter plot in terms of a context, and add points to a scatter plot given information about an individual in the population. They compare individuals represented by different points and informally discuss trends in the data.

There are two levels of analysis needed to successfully make sense of scatter plots: what is happening for a particular individual, and what is happening at a global level for the entire population. The ability to move between these two zoom levels develops over time. In this lesson, students spend a lot of time looking at the details of a scatter plot, naming the quantities represented in a scatter plot, and focusing on the meaning of individual points (MP6).

Alignments

Building On

- 5.G.A.2: Represent real world and mathematical problems by graphing points in the first quadrant of the coordinate plane, and interpret coordinate values of points in the context of the situation.

Addressing

- 8.SP.A.1: 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.
- 8.SP.A.3: Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope 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

- MLR1: Stronger and Clearer Each Time

- MLR5: Co-Craft Questions
- Think Pair Share

Student Learning Goals

Let's investigate points in scatter plots.

3.1 The Giant Panda

Warm Up: 5 minutes

The purpose of this warm-up is for students to interpret the meaning of a single point in a scatter plot by looking at the point's coordinates and the graph's axis labels (MP2).

Building On

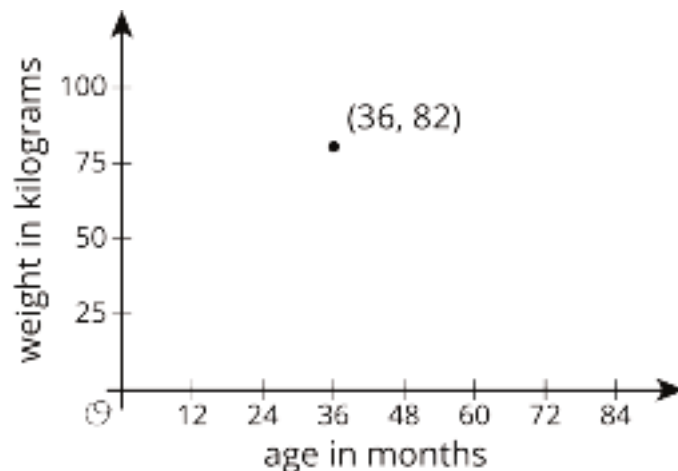
- 5.G.A.2

Launch

Display the graph for all to see. Give students 1 minute of quiet think time followed by a whole-class discussion.

Student Task Statement

A giant panda lives in a zoo. What does the point on the graph tell you about the panda?



Student Response

At 36 months of age, the panda weighed 82 kilograms.

Activity Synthesis

Ask students to share their responses. Record and display their responses for all to see. If all of the students responds, "At 36 months of age, the panda weighed 82 kilograms," ask them the following questions to clarify where they found that information in the graph:

- "What did the first coordinate of the point tell you?" (The first coordinate shows how far to go on the horizontal axis.)
- "What did the second coordinate of the point tell you?" (The second coordinate shows how far to go on the vertical axis.)
- "How did you know the meaning of each of the numbers?" (You can read the meaning of the numbers by looking at the labels on the axes.)
- "What specific information did each axis give you?" (The labels on the axes indicate attribute, age or weight, and the units in which each attribute is measured, months or kilograms.)

3.2 Weight and Fuel Efficiency

15 minutes (there is a digital version of this activity)

The analysis of scatter plots continues with data about the mass of automobiles and their fuel efficiency. Again, students connect points in the scatter plot with rows of data in a table, but now there is a third column that gives a name to each pair.

After picking out information about points in the scatter plot and table, students make their first foray into thinking about how scatter plots can help them make predictions. Specifically, when looking at two particular points in the scatter plot, students are asked if the results are surprising given the overall trend of the data. In later lessons, students develop more sophisticated tools for answering questions like this (specifically, by fitting lines to the data), so a heuristic discussion is all that is expected at this point.

Addressing

- 8.SP.A.1

Instructional Routines

- MLR5: Co-Craft Questions

Launch

Explain a bit about fuel efficiency so that students understand this measurement. "Fuel efficiency" is a measure of the average distance a car will travel using a certain amount of gas. Commonly this is measured in "miles per gallon." For example, a car that has a fuel efficiency of 25 miles per gallon (mpg) should be able to drive approximately 25 miles while using up a gallon of gas. Many factors can influence fuel efficiency, including the way an engine is engineered (to produce more power, an engine may use more gasoline), driving conditions (more frequent stopping and starting or hills will

use fuel less efficiently), and what accessories are being used (air conditioning requires energy that is not used for actually driving the car).

Before beginning the activity, ask “What would a scatter plot including these two variables look like?” to encourage students to predict how a car’s mass might influence fuel efficiency.

If using the digital version of the activity, after completing the launch discussion, students will use an applet to analyze a scatter plot, identify points, and draw conclusions about the data. The digital activity mirrors the analog activity.

Access for English Language Learners

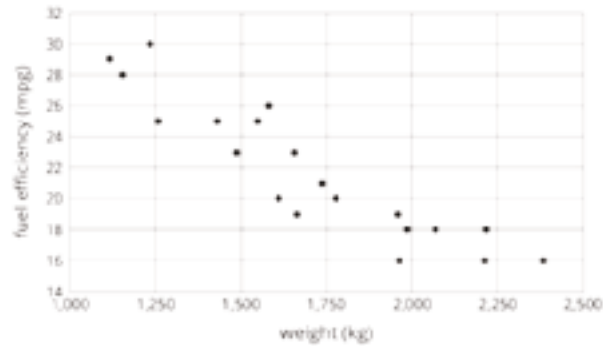
Writing, Conversing: MLR5 Co-Craft Questions. Display only the table without revealing the graph or questions that follow. Ask students to work with a partner to write possible questions that could be answered by the data in the table. Invite 2–3 groups to share their questions with the class. Look to see whether/how the weights and fuel efficiencies of the vehicles shows up in students’ questions. Next, reveal the questions of the activity. This routine helps students consider the context of this problem and to increase awareness about language used to talk about weights and fuel efficiencies.

Design Principle(s): Maximize meta-awareness; Support sense-making

Student Task Statement

The table and scatter plot show weights and fuel efficiencies of 18 cars.

car	weight (kg)	fuel efficiency
A	1,549	25
B	1,610	20
C	1,737	21
D	1,777	20
E	1,486	23
F	1,962	16
G	2,384	16
H	1,957	19
I	2,212	16
J	1,115	29
K	2,068	18
L	1,663	19
M	2,216	18
N	1,432	25
O	1,987	18
P	1,580	26
Q	1,234	30
R	1,656	23



1. Which point in the scatter plot represents Car L's measurements?
2. What is the fuel efficiency of the car with the greatest weight?
3. What is the weight of the car with the greatest fuel efficiency?
4. Car S weighs 1,912 kilograms and gets 16 miles per gallon. On the scatter plot, plot a point that represents Car S's measurements.
5. Cars N and O, shown in the scatter plot, are made by the same company. Compare their weights and fuel efficiencies. Does anything surprise you about these cars?
6. A different company makes Cars F and G. Compare their weights and fuel efficiencies. Does anything surprise you about these cars?

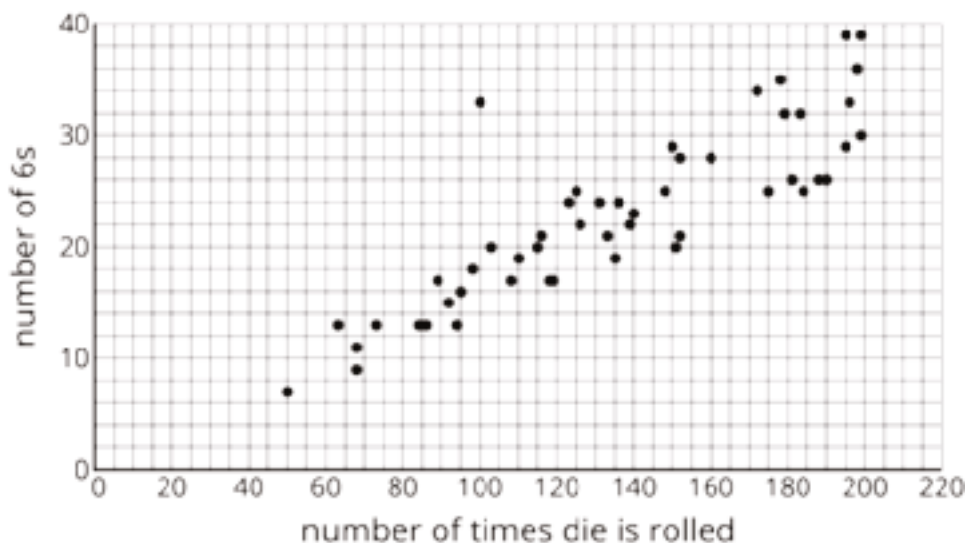
Student Response

1. The point with coordinates (1,663, 19)
2. 16 mpg
3. 1,234 kg
4. The point with coordinates (1,912, 16)

5. Car N has a lower weight and a higher fuel efficiency than car O. Answers vary. Sample response: Nothing is surprising about these cars because they fit in with the trend of the data.
6. Their weights are quite different, but they still have the same fuel efficiency. Answers vary. Sample response: It is surprising that cars with such different weights made by the same company have the same fuel efficiency since generally the heavier the car, the less fuel efficient it is.

Are You Ready for More?

After a board game competition, the tournament director collects 50 dice from the games played and rolls each one until he gets bored and tries a different one. The scatter plot shows the number of times he rolled each die and the number of 6s that resulted during those rolls.



Select a point in the scatter plot and give its approximate coordinates, then tell the story of that point in the context of the problem.

Student Response

Answers vary. Sample response: The point at (100, 33) lies apart from the rest of the data. Someone may have brought a die from home that is weighted to roll more 6s, expecting to cheat and win the tournament.

Activity Synthesis

The purpose of the discussion is for students to compare pairs of data in both representations and to begin thinking about trends in the data. Ask students to share their thinking about the last two questions. Draw out the fact that cars N and O seem to fit the overall trend that an increase in mass typically corresponds to a decrease in fuel efficiency but cars F and G don't fit that trend when compared with each other. Mention that we will see many examples of trends that don't necessarily apply to individuals in a population.

Consider asking some of the following questions to conclude the discussion:

- “How did you find the automobile with the greatest weight?” (I looked for the point farthest to the right.)
- “How did you find the automobile with the greatest fuel efficiency?” (I looked for the point that is highest.)
- “The data in this table is from automobiles that run solely on gasoline. Where do you think a hybrid car would appear on the graph?” (A hybrid car would use less gas for the same weight, so its point would be below the other points with a similar weight.)
- “Did the scatter plot match your prediction?”

3.3 Coat Sales

15 minutes

In this activity, students continue to identify points on a scatter plot as representatives of a single month when two variables are measured. Students are also asked to explain the meaning of some abstract points in the context of the problem and identify when it does not make sense to extrapolate information from the graph based on the context (MP2).

Monitor for students who understand that 60 degrees Celsius is unreasonable for an average monthly temperature in most parts of the world, as well as those who identify the trend to predict negative sales which does not make sense in this context either.

Addressing

- 8.SP.A.1
- 8.SP.A.3

Instructional Routines

- MLR1: Stronger and Clearer Each Time
- Think Pair Share

Launch

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

Ask students to predict what a scatter plot might look like when the temperature is along the x -axis and the sales of coats at a store is along the y -axis.

Access for Students with Disabilities

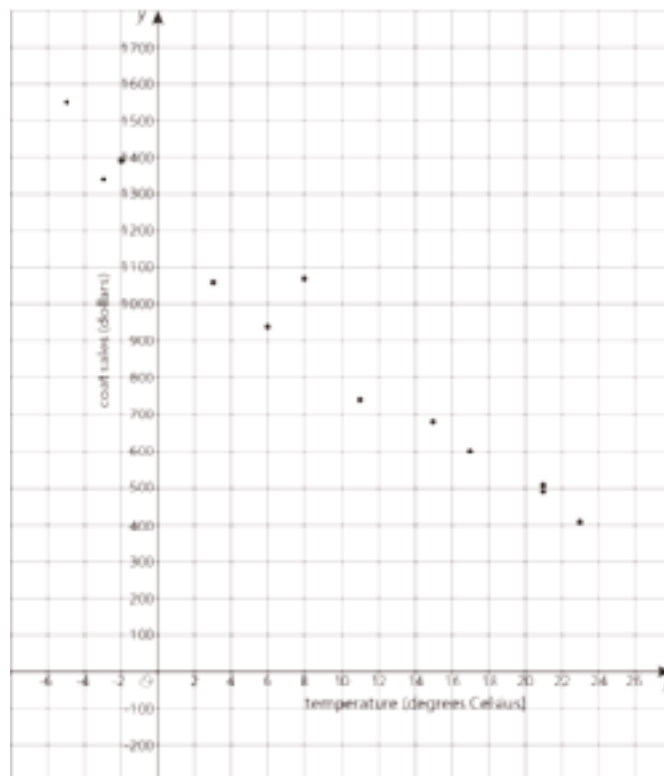
Representation: Internalize Comprehension. Demonstrate and encourage students to use color coding and annotations to highlight connections between representations in a problem. For example, ask students to use the different colors to represent and annotate temperature and coat sales in the graph and table.

Supports accessibility for: Visual-spatial processing

Student Task Statement

A clothing store keeps track of the average monthly temperature in degrees Celsius and coat sales in dollars.

temperature (degrees Celsius)	coat sales (dollars)
-5	1,550
-3	1,340
3	1,060
8	1,070
15	680
21	490
23	410
21	510
17	600
11	740
6	940
-2	1,390



1. What does the point (15, 680) represent?

2. For the month with the lowest average temperature, estimate the total amount made from coat sales. Explain how you used the table to find this information.
3. For the month with the smallest coat sales, estimate the average monthly temperature. Explain how you used the scatter plot to find this information.
4. If there were a point at $(0, A)$ what would it represent? Use the scatter plot to estimate a value for A .
5. What would a point at $(B, 0)$ represent? Use the scatter plot to estimate a value for B .
6. Would it make sense to use this trend to estimate the value of sales when the average monthly temperature is 60 degrees Celsius? Explain your reasoning.

Student Response

1. When the average monthly temperature is 15 degree Celsius, the store sold \$680 of coats.
2. \$1,550. In the table, I found the lowest temperature and wrote the value for sales in the same row.
3. 23 degrees Celsius. In the graph, I found the lowest point and estimated the x coordinate.
4. When the temperature averaged 0 degrees Celsius for the month, the store sold A dollars of coats. Estimates vary. Correct answers should range between \$1,200 and \$1,300.
5. When the store sold no coats for the month, the average monthly temperature is B degrees Celsius. Estimates vary. Correct answers should be approximately 30 degrees Celsius.
6. It does not make sense. 60 degrees Celsius (140 Fahrenheit) is hotter than has been recorded on Earth. Also, based on the trend, there would be negative dollars sold which does not really make sense.

Activity Synthesis

The purpose of the discussion is for students to understand that each point in the graph represents a single month in which two measurements are made. Select students to share their responses, including those identified for the last problem.

Consider asking some of the following questions:

- “When discussing functions, what is a point of the form $(0, A)$ called? What about a point of the form $(B, 0)$?” (The y -intercept. The x -intercept.)
- “In a scatter plot is there a y -intercept?” (It depends on the context. For some situations, it may not make sense for the variable represented along the x -axis to be zero. For example if the variable is the height of a building. On other situations, there may be more than one point along the y -axis. For example, in a situation like the one in the task, there may be multiple months where the average temperature is 0 degree Celsius.)

- "When the average monthly temperature goes up, what seems to be happening to sales?" (They are decreasing.)
-

Access for English Language Learners

Writing, Speaking: MLR1 Stronger and Clearer Each Time. Use this routine to give students a structured opportunity to revise and refine 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., "Can you give an example?", "Why do you think...?", "How do you know...?", etc.). Students can borrow ideas and language from each partner to strengthen their final version.

Design Principle(s): Optimize output (for explanation)

Lesson Synthesis

Revisit the idea that a point in a scatter plot represents an individual in a population. Ask students:

- "What kind of information does a point in a scatter plot represent?" (Two measurements for an individual in a population.)
- "How do we know what information is captured by a point in a scatter plot?" (The axis labels tell us how to interpret the coordinates of the points.)
- "Do trends in a scatter plot necessarily apply to individuals in the population?" (No.)

3.4 Quarterbacks

Cool Down: 5 minutes

Students continue to practice comparing data from a table and scatter plot by identifying points in the scatter plot and their corresponding entries in a table. Additionally, students show their understanding of the connection by plotting an additional point to add to the data shown.

Addressing

- 8.SP.A.1

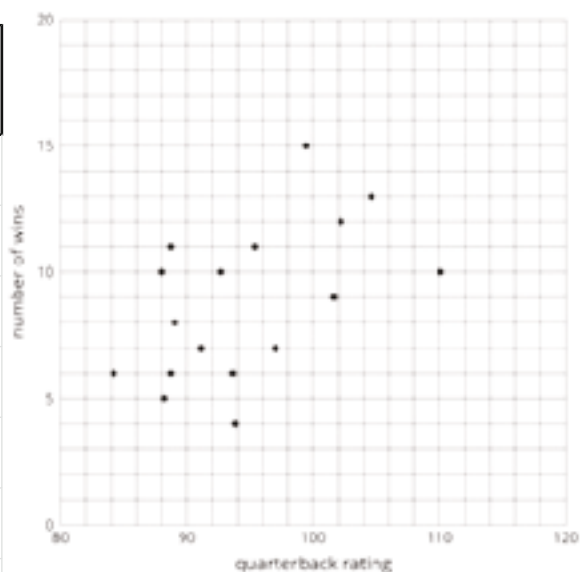
Launch

Explain that quarterback rating is a way of rating a quarterback in football on a scale of 0 to 158.3 based on many variables (such as pass attempts, interceptions, and touchdown passes). A higher rating generally represents a quarterback with better passing statistics. When the quarterback starts a game, he is playing from the beginning of the game and is usually the best option available for the team at that position. The data from this activity refers to quarterbacks who were able to start in at least 16 games for their team in a season.

Student Task Statement

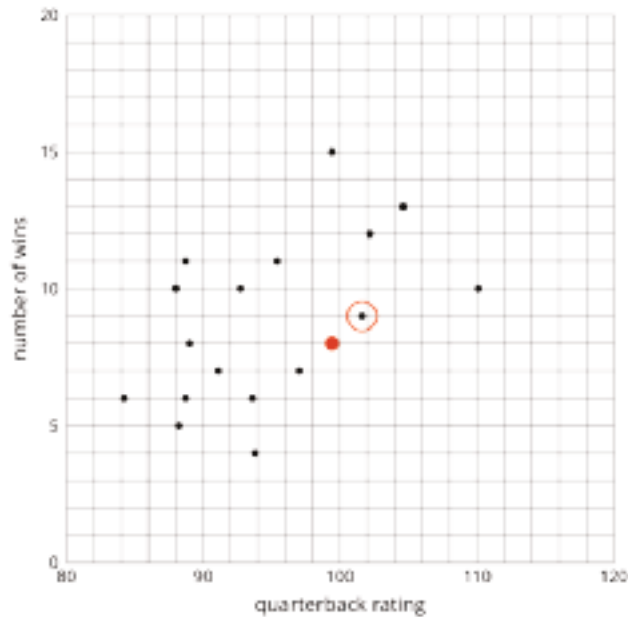
Here are a table and scatter plot that show ratings and wins for quarterbacks who started 16 games this season.

player	quarterback rating	number of wins
A	93.8	4
B	102.2	12
C	93.6	6
D	89	8
E	88.2	5
F	97	7
G	88.7	6
H	91.1	7
I	92.7	10
J	88	10
K	101.6	9
L	104.6	13
M	84.2	6
N	99.4	15
O	110.1	10
P	95.4	11
Q	88.7	11



1. Circle the point in the scatter plot that represents Player K's data.
2. Which quarterback's data are represented by the point farthest to the left?
3. Player R is not included in the table because he did not start 16 games this year. He did have a quarterback rating of 99.4 and his team won 8 games. On the scatter plot, plot a point that represents Player R's data.

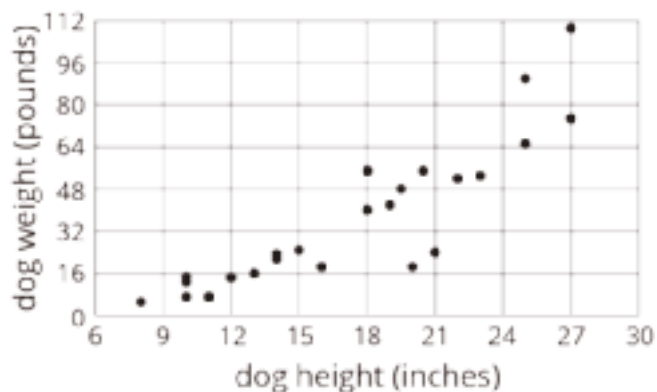
Student Response



1. The circled point on the scatter plot
2. Player M
3. The added point to the scatter plot, plotted larger for visibility

Student Lesson Summary

Scatter plots show two measurements for each individual from a group. For example, this scatter plot shows the weight and height for each dog from a group of 25 dogs.



We can see that the tallest dogs are 27 inches, and that one of those tallest dogs weighs about 75 pounds while the other weighs about 110 pounds. This shows us that dog weight is not a function of dog height because there would be two different outputs for the same input. But we can see a general trend: Taller dogs tend to weigh more than shorter dogs. There are exceptions. For example, there is a dog that is 18 inches tall and weighs over 50 pounds, and there is another dog that is 21 inches tall but weighs less than 30 pounds.



When we collect data by measuring attributes like height, weight, area, or volume, we call the data *numerical data* (or measurement data), and we say that height, weight, area, or volume is a *numerical variable*. Upcoming lessons will discuss how to identify and describe trends in data that has been collected.

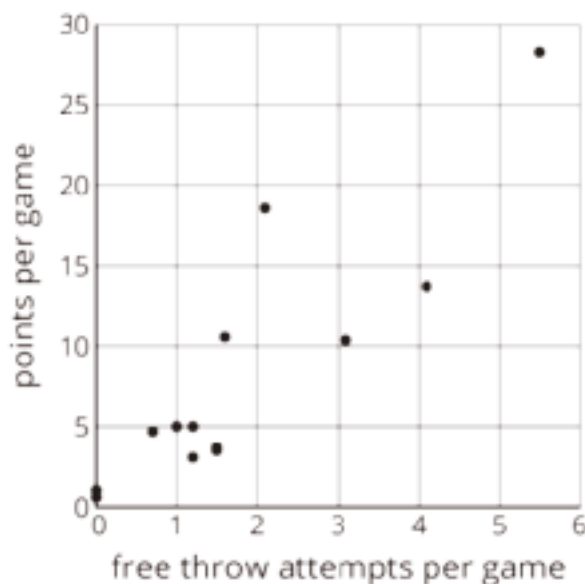
Lesson 3 Practice Problems

Problem 1

Statement

Here is a table and a scatter plot that compares points per game to free throw attempts for a basketball team during a tournament.

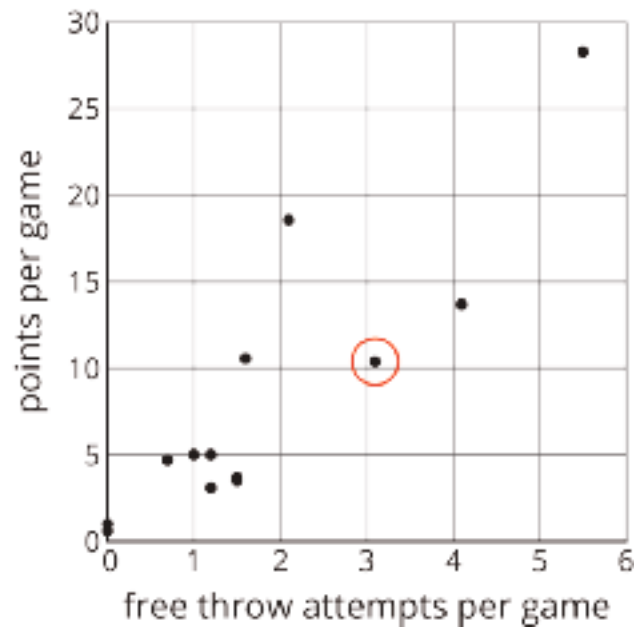
player	free throw attempts	points
player A	5.5	28.3
player B	2.1	18.6
player C	4.1	13.7
player D	1.6	10.6
player E	3.1	10.4
player F	1	5
player G	1.2	5
player H	0.7	4.7
player I	1.5	3.7
player J	1.5	3.5
player K	1.2	3.1
player L	0	1
player M	0	0.8
player N	0	0.6



- Circle the point that represents the data for Player E.
- What does the point (2.1, 18.6) represent?
- In that same tournament, Player O on another team scored 14.3 points per game with 4.8 free throw attempts per game. Plot a point on the graph that shows this information.

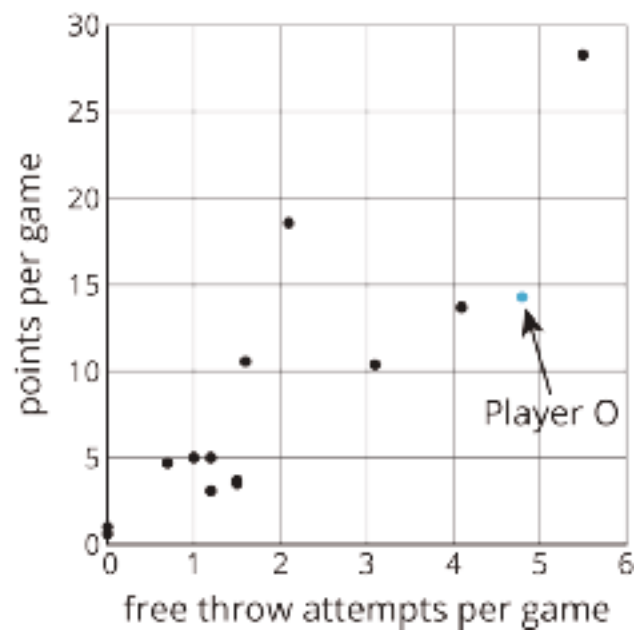
Solution

a.



b. The point (2.1, 18.6) represents the free throw attempts and points per game for Player B.

c.



Problem 2

Statement

Select **all** the representations that are appropriate for comparing exam score to number of hours of sleep the night before the exam.

- A. Histogram
- B. Scatter plot
- C. Dot plot
- D. Table
- E. Box plot

Solution

["B", "D"]

(From Unit 6, Lesson 2.)

Problem 3

Statement

A cylinder has a volume of $36\pi \text{ cm}^3$ and height h . Complete this table for the volume of other cylinders with the same radius but different heights.

height (cm)	volume (cm^3)
h	36π
$2h$	
$5h$	
$\frac{h}{2}$	
$\frac{h}{5}$	

Solution

height (cm)	volume (cm ³)
h	36π
$2h$	72π
$5h$	180π
$\frac{h}{2}$	18π
$\frac{h}{5}$	$\frac{36}{5}\pi$

(From Unit 5, Lesson 17.)

Lesson 4: Fitting a Line to Data

Goals

- Compare and contrast (orally) values in a data set with predictions made using a given line.
- Comprehend that a model of data, such as a line of fit, can be used to predict values that are not given in the data.
- Identify (orally) obvious outliers on a scatter plot.

Learning Targets

- I can pick out outliers on a scatter plot.
- I can use a model to predict values for data.

Lesson Narrative

In the previous lesson, students focused primarily on the details of a scatter plot. In this lesson, their focus becomes more holistic and they begin to see a set of data points as a single thing that can be analyzed, not just a bunch of disconnected points. For the first time students see that sometimes we can model the relationship between two variables with a line (MP4), although they continue to analyze the connections between the scatter plot and the line by comparing individual points. As they zoom out and see these connections, students begin to see the structure of the scatter plot (MP7) and to use that structure to reason abstractly and quantitatively (MP2).

Alignments

Addressing

- 8.SP.A.1: 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.
- 8.SP.A.2: 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.

Instructional Routines

- MLR8: Discussion Supports
- Notice and Wonder
- Think Pair Share

Student Learning Goals

Let's look at the scatter plots as a whole.

4.1 Predict This

Warm Up: 5 minutes

In previous lessons, students were asked to interpret individual points in a scatter plot or compare two points with each other. In this warm-up, students are asked think about a point in the context of the entire data set. They compare two different possible predictions for the dependent variable, given a value for the independent variable. This sets them up to understand and interpret a linear model for the relationship between independent and dependent variables. They use the potential location of a point not included in the scatter plot to answer a question about the context, based on the visual structure of the scatter plot (MP7, MP2).

Addressing

- 8.SP.A.1

Instructional Routines

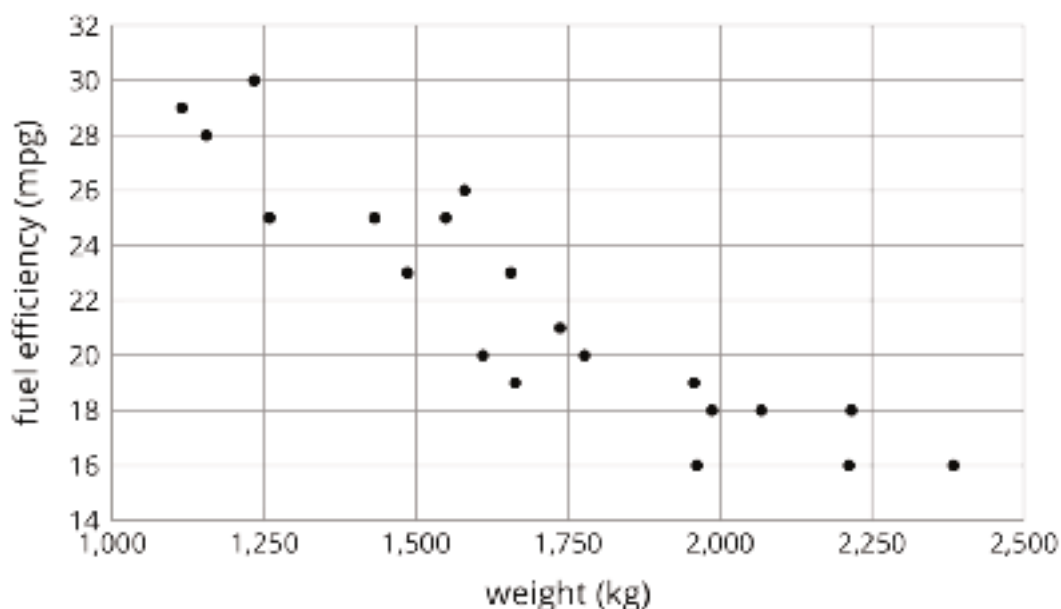
- Think Pair Share

Launch

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

Student Task Statement

Here is a scatter plot that shows weights and fuel efficiencies of 20 different types of cars.



If a car weighs 1,750 kg, would you expect its fuel efficiency to be closer to 22 mpg or to 28 mpg? Explain your reasoning.

Student Response

We expect the fuel efficiency to be closer to 22 mpg. There are several cars close to that weight, and their fuel efficiency is between 18 and 22 mpg. The cars that have a fuel efficiency close to 28 mpg have a weight less than 1,250 kg, although there is one car that has a weight greater than 1,500 kg and a fuel efficiency of 26 mpg. But that one has a higher fuel efficiency than any other car with a weight between 1,500 and 1,750 kg.

Activity Synthesis

Display the graph for all to see. Poll the class to see if they think the fuel efficiency is closer to 22 mpg or 28 mpg. If they are all in agreement that the answer is closer to 22 mpg, ask a few students to share their reasoning. If there is disagreement, ask students to share their reasoning and come to an agreement. If it does not come up in the discussion, ask students to look at cars whose fuel efficiency is close to 28 mpg and note that their weights are quite a bit less. Then look at cars with a weight close to 1,750 kg, and note that their fuel efficiency is between 18 and 22 mpg. As a whole class, decide where to plot both potential points, and point out that one is close to the other nearby values and one is very far away.

4.2 Shine Bright

15 minutes (there is a digital version of this activity)

In this lesson, the meaning of the words *model* and *modeling* are explained in terms of a linear model. Students are not expected to define the words, but should be comfortable understanding and using them. A model is used to predict prices for diamonds not included in the data as well as compare existing data points to the model.

Addressing

- 8.SP.A.1
- 8.SP.A.2

Instructional Routines

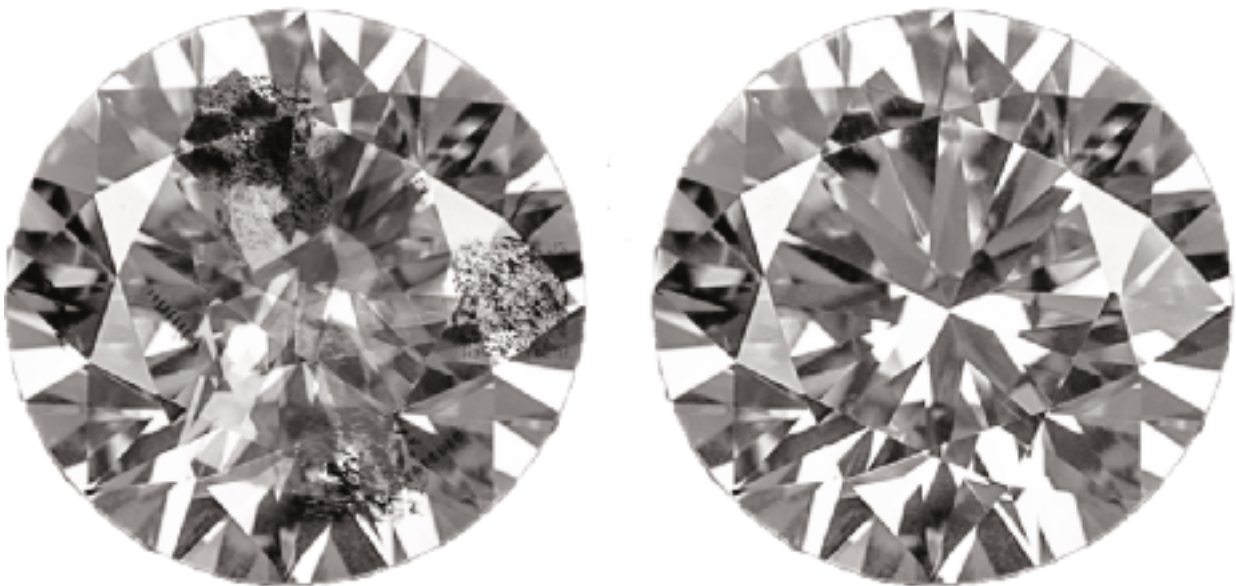
- MLR8: Discussion Supports
- Think Pair Share

Launch

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



Show the first image for all to see. Ask students, “Which diamond do you think has the highest price? Which diamond do you think has the lowest price?” Most likely, they will guess that the largest has the highest price and the smallest has the lowest price. Validate this intuition and let them know that the size of a diamond is *one* thing that its price is based on. Tell students that the size of a diamond is commonly measured in *carats*, which is a measure of weight. (One carat is equivalent to 200 milligrams.)



Show the second image for all to see. Ask students, “Which diamond do you think has the higher price?” Poll the class, and ask a few students to share their reasoning. Tell students that these diamonds are the same size, but they are not the same price. The price of a diamond is not just based on size, but also on how many flaws are inside the stone. Since the diamond on the right has fewer flaws (which makes it more sparkly), it has a higher price than the diamond on the left.

Tell students, “You will look at some data for the prices of diamonds of different sizes. You have used mathematics to analyze real-world situations, identifying variables in a situation and describing their relationships mathematically. This process is called *modeling*, and the mathematical description is called a *model*. Sometimes you made assumptions about the situation or ignored some features so that the model would be simpler.”

Access for Students with Disabilities

Action and Expression: Develop Expression and Communication. Invite students to talk about their ideas with a partner before writing them down. Display sentence frames to support students when they explain their ideas. For example, “It looks like...” and “We are trying to...”

Supports accessibility for: Language; Organization

Access for English Language Learners

Speaking: MLR8 Discussion Supports. Provide sentence frames to support students as they work with their partner and explain their reasoning. For example, “I think ____ because ____.” or “I (agree/disagree) because ____.”

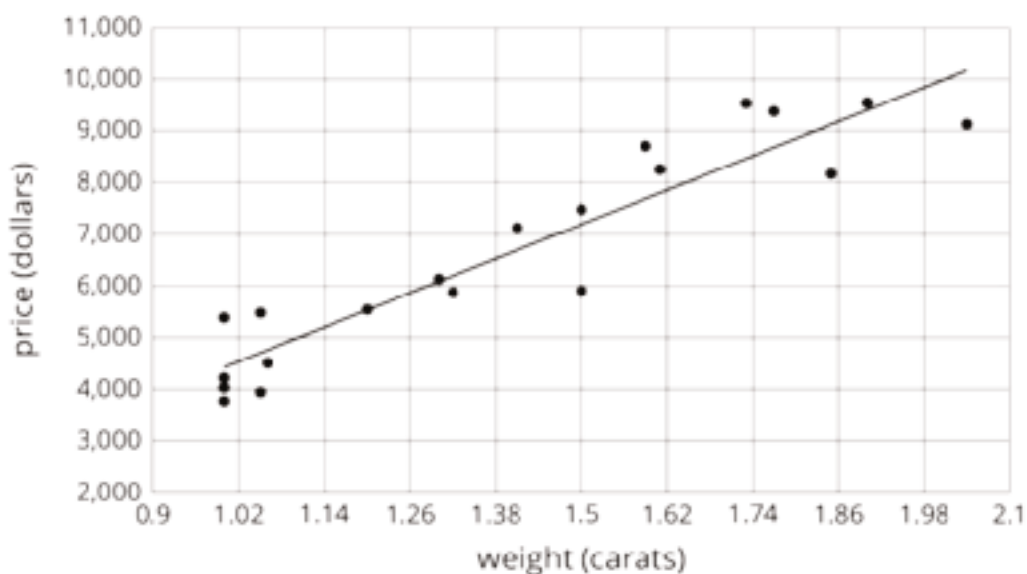
Design Principle(s): Support sense-making; Optimize output for (explanation)

Student Task Statement

Here is a table that shows weights and prices of 20 different diamonds.

weight (carats)	actual price (dollars)	predicted price (dollars)
1	3,772	4,429
1	4,221	4,429
1	4,032	4,429
1	5,385	4,429
1.05	3,942	4,705
1.05	4,480	4,705
1.06	4,511	4,760
1.2	5,544	5,533
1.3	6,131	6,085
1.32	5,872	6,195
1.41	7,122	6,692
1.5	7,474	7,189
1.5	5,904	7,189
1.59	8,706	7,686
1.61	8,252	7,796
1.73	9,530	8,459
1.77	9,374	8,679
1.85	8,169	9,121
1.9	9,541	9,397
2.04	9,125	10,170

The scatter plot shows the prices and weights of the 20 diamonds together with the graph of $y = 5,520x - 1,091$.



The function described by the equation $y = 5,520x - 1,091$ is a *model* of the relationship between a diamond's weight and its price.

This model *predicts* the price of a diamond from its weight. These predicted prices are shown in the third column of the table.

1. Two diamonds that both weigh 1.5 carats have different prices. What are their prices? How can you see this in the table? How can you see this in the graph?
2. The model predicts that when the weight is 1.5 carats, the price will be \$7,189. How can you see this in the graph? How can you see this using the equation?
3. One of the diamonds weighs 1.9 carats. What does the model predict for its price? How does that compare to the actual price?
4. Find a diamond for which the model makes a very good prediction of the actual price. How can you see this in the table? In the graph?
5. Find a diamond for which the model's prediction is not very close to the actual price. How can you see this in the table? In the graph?

Student Response

1. \$7,474 and \$5,904. You can see this in the two rows in the table that show 1.5 in the first column. You can see this in the graph by finding the two points with $x = 1.5$.
2. The point $(1.5, 7,189)$ is on the graph of the line. We can substitute $x = 1.5$ into the equation to get $y = 5,520 \cdot 1.5 - 1,091 = 7,189$.
3. The model predicts that the price will be \$9,397, because $y = 5,520 \cdot 1.9 - 1,091 = 9,397$. The actual price is \$9,541. The actual price is \$144 greater than the prediction, because $9,541 - 9,397 = 144$.

4. The best prediction is made for the diamond with a weight of 1.2 carats—the prediction is only off by \$11. You can see this in the table in the row with 1.2 in the first column. You can see that the corresponding point in the scatter plot is very close to the line.
5. Answers vary. Sample response: The prediction for one of the diamonds that weighs 1.5 carats is off by over \$1,000. You can see this in the second row with a 1.5 in the first column. You can see that the corresponding point in the scatter plot is far below the line.

Activity Synthesis

The goal of this discussion is to help students understand the relationship between the data and a linear model of the data.

To highlight there relationship, ask:

- "What does a point in the scatter plot represent?" (The *actual* weight and price of a diamond.)
- "What does the line represent?" (The *predicted* price of a diamond based on its weight.)
- "What does it mean when a point is close to the line? When it is far away from the line?" (When it is close, the model predicts the price well. The farther away in the vertical direction, the worse the prediction.)
- "How can you use the graph to predict the price of a diamond that weighs 1.1 carats? How can you use the equation?" (Find the point on the line that lines up with $x = 1.1$ or substitute $x = 1.1$ into the model equation.)

4.3 The Agony of the Feet

15 minutes (there is a digital version of this activity)

A scatter plot is shown and the points interpreted in context. Later, a linear model is graphed with the scatter plot to help students see an obvious outlier (MP7). In the discussion, the term outlier is introduced.

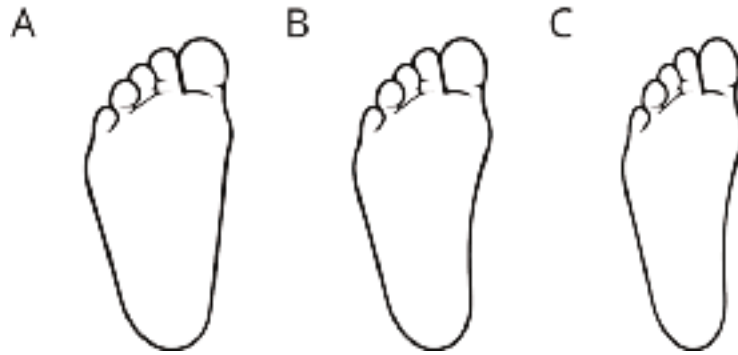
Addressing

- 8.SP.A.1
- 8.SP.A.2

Instructional Routines

- MLR8: Discussion Supports
- Notice and Wonder

Launch



Ask students, “What do you notice? What do you wonder?” Ideally, they will notice that these are pictures of a foot of three different people. All three feet are approximately the same length but different widths. Tell students that all of these feet are a size 8. However, they wouldn’t all necessarily find the same shoe equally comfortably, because of the varying widths. (Besides the numerical size, some shoes also come in different widths.) Students should understand that human feet can vary in both length and width.

If using the digital activity, after this discussion, clarify for students that a scatter plot has been created for them to analyze feet dimensions, using technology. Students can then work to complete the digital task which includes using a line of best fit to identify outliers.

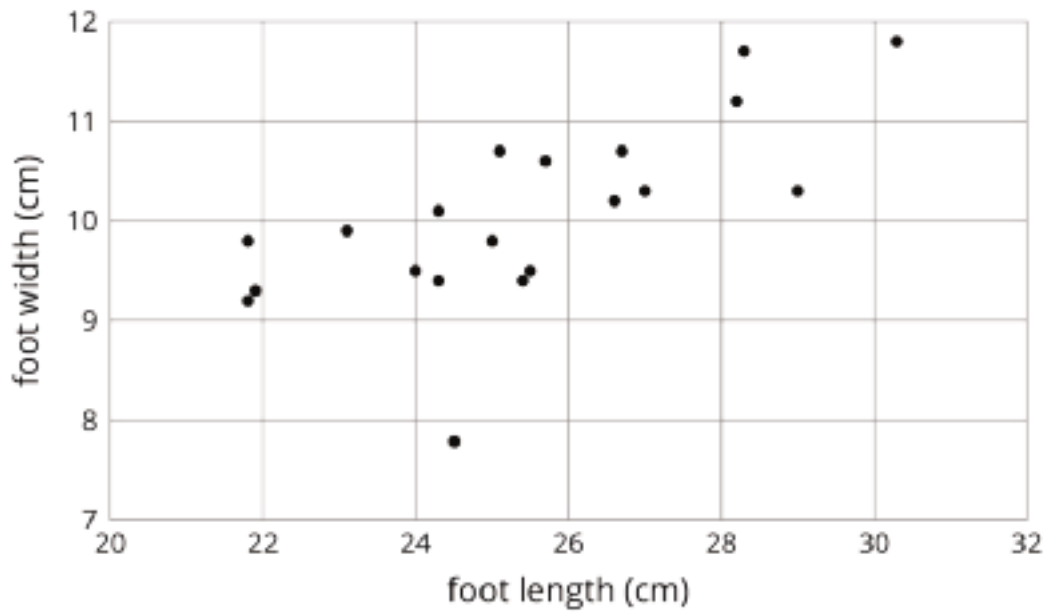
Access for English Language Learners

Speaking: MLR8 Discussion Supports. Use this routine to clarify the meaning of “width” and “length” in the context of the task. Connect the terms width and length multi-modally by utilizing different types of sensory inputs, such as demonstrating the width and length on the images of each foot using gestures or inviting students to do so. Invite students to chorally repeat phrases that include these words in context. This will help students produce and make sense of the language needed to understand the context of the task.

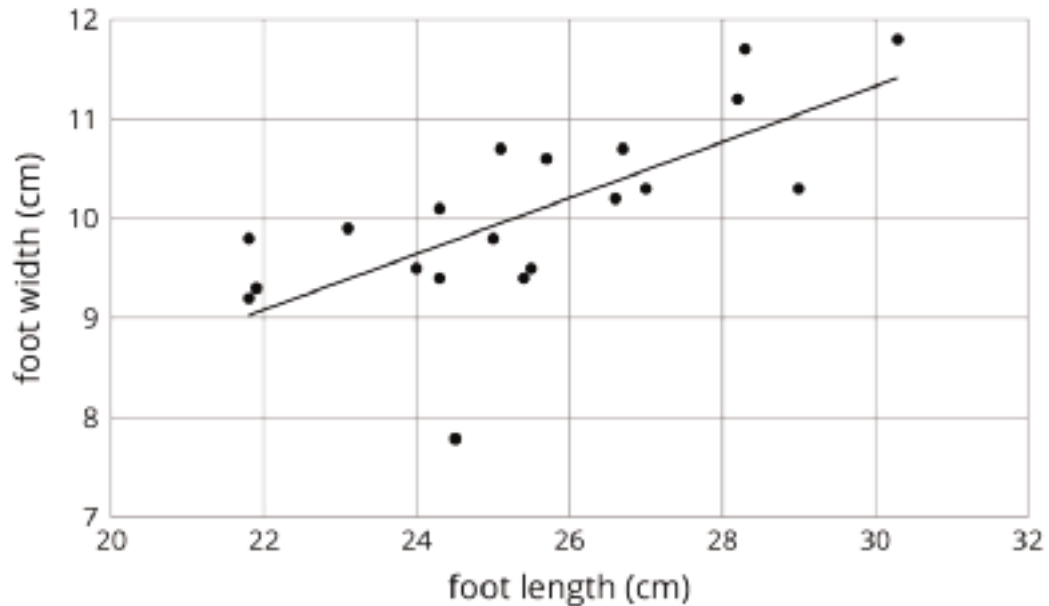
Design Principle(s): Support sense-making

Student Task Statement

Here is a scatter plot that shows lengths and widths of 20 different left feet.



1. Estimate the widths of the longest foot and the shortest foot.
2. Estimate the lengths of the widest foot and the narrowest foot.
3. Here is the same scatter plot together with the graph of a model for the relationship between foot length and width.



Circle the data point that seems weird when compared to the model. What length and width does that point represent?

Student Response

1. Longest has a width of about 11.9 cm, shortest has a width of about 9.3 cm

2. Widest has a length of about 30.4 cm, narrowest has a length of about 24.5 cm
3. The point that “seems weird” represents length approximately 24.5 cm and width about 7.8 cm

Activity Synthesis

Introduce the term **outlier**. An outlier is a point that is separated from the rest of the data. Sometimes data sets have outliers. Sometimes that’s because there really is a data point that is very different than the others. Sometimes it is because there was an error in collecting the data. Sometimes it is because there was an error in entering the data. When there are outliers, one has to make a judgement about whether to include it in the analysis or not.

Lesson Synthesis

To help students see the connection between a scatter plot and a linear model for a data set, ask:

- "What kind of *model* for a data set did we investigate today?" (A linear model.)
- "What does this kind of model help us do?" (See the trend in the data more clearly and make predictions.)
- "What does it mean when a data point is closer in the vertical direction to the line that represents a linear model? What does it mean when a data point is farther from the line in the vertical direction?" (When the data point is closer in the vertical direction to the line, it represents data that fits the prediction well. When the point is farther from the line, the data does not fit the prediction well.)
- "In your own words, how can you identify an outlier from a scatter plot?" (A point that is far from the other points in the scatter plot represents an outlier.)

4.4 A 1 Foot Foot

Cool Down: 5 minutes

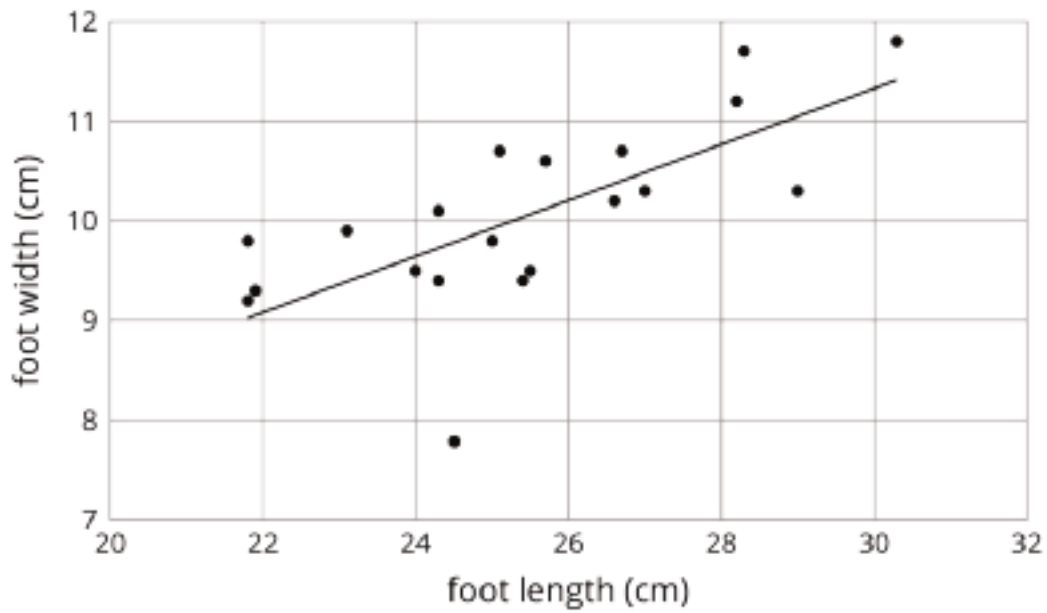
Given a scatter plot, students find the point that is closest to a given value and then compare this point to the predicted value based on a linear model.

Addressing

- 8.SP.A.1
- 8.SP.A.2

Student Task Statement

Here is a scatter plot that shows lengths and widths of 20 left feet, together with the graph of a model of the relationship between foot length and width.



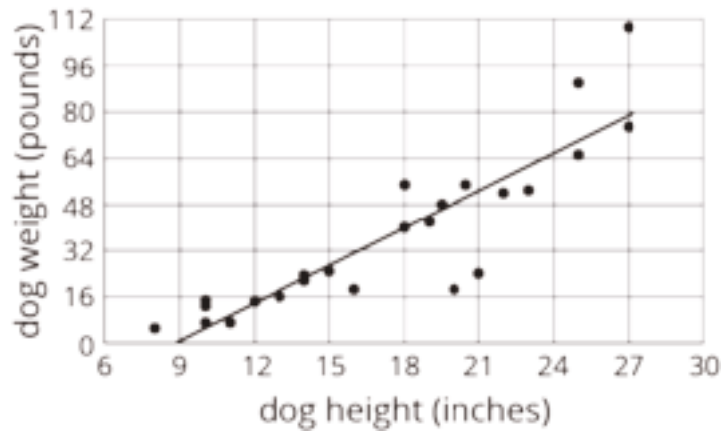
1. Draw a box around the point that represents the foot with length closest to 29 cm.
2. What is the approximate width of this foot?
3. What width does the model predict for a foot with length 29 cm?

Student Response

1. A box is drawn around the point at approximately (29.1, 10.4).
2. About 10.4 cm
3. About 11 cm (or a little more than 11 cm)

Student Lesson Summary

Sometimes, we can use a linear function as a model of the relationship between two variables. For example, here is a scatter plot that shows heights and weights of 25 dogs together with the graph of a linear function which is a model for the relationship between a dog's height and its weight.



We can see that the model does a good job of predicting the weight given the height for some dogs. These correspond to points on or near the line. The model doesn't do a very good job of predicting the weight given the height for the dogs whose points are far from the line.

For example, there is a dog that is about 20 inches tall and weighs a little more than 16 pounds. The model predicts that the weight would be about 48 pounds. We say that the model *overpredicts* the weight of this dog. There is also a dog that is 27 inches tall and weighs about 110 pounds. The model predicts that its weight will be a little less than 80 pounds. We say the model *underpredicts* the weight of this dog.

Sometimes a data point is far away from the other points or doesn't fit a trend that all the other points fit. We call these **outliers**.

Glossary

- outlier

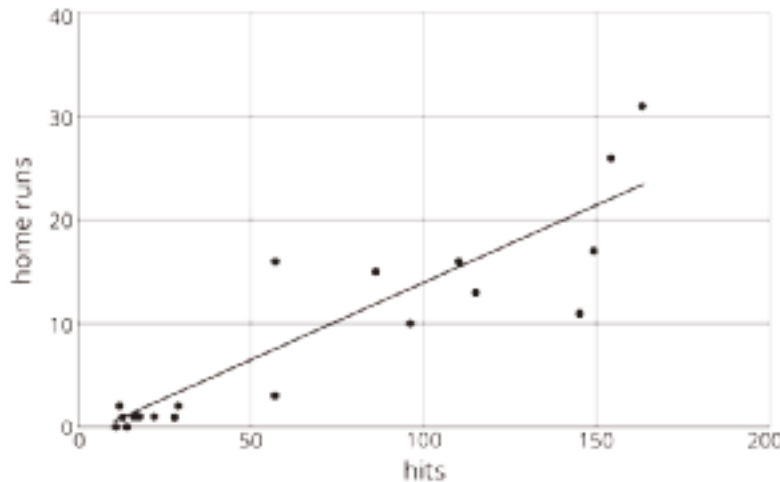
Lesson 4 Practice Problems

Problem 1

Statement

The scatter plot shows the number of hits and home runs for 20 baseball players who had at least 10 hits last season. The table shows the values for 15 of those players.

The model, represented by $y = 0.15x - 1.5$, is graphed with a scatter plot.



hits	home runs	predicted home runs
12	2	0.3
22	1	1.8
154	26	21.6
145	11	20.3
110	16	15
57	3	7.1
149	17	20.9
29	2	2.9
13	1	0.5
18	1	1.2
86	15	11.4
163	31	23
115	13	15.8
57	16	7.1
96	10	12.9

Use the graph and the table to answer the questions.

- Player A had 154 hits in 2015. How many home runs did he have? How many was he predicted to have?
- Player B was the player who most outperformed the prediction. How many hits did Player B have last season?
- What would you expect to see in the graph for a player who hit many fewer home runs than the model predicted?

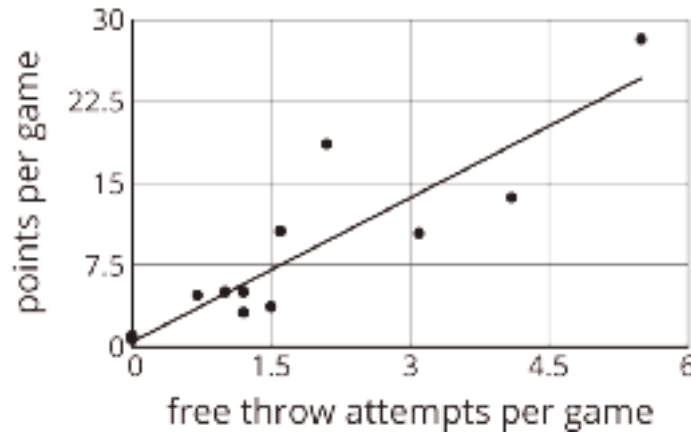
Solution

- Home runs: 26. Predicted home runs: 21.6
- 57
- The point should be much lower on the graph than the line.

Problem 2

Statement

Here is a scatter plot that compares points per game to free throw attempts per game for basketball players in a tournament. The model, represented by $y = 4.413x + 0.377$, is graphed with the scatter plot. Here, x represents free throw attempts per game, and y represents points per game.



- Circle any data points that appear to be outliers.
- What does it mean for a point to be far above the line in this situation?
- Based on the model, how many points per game would you expect a player who attempts 4.5 free throws per game to have? Round your answer to the nearest tenth of a point per game.
- One of the players scored 13.7 points per game with 4.1 free throw attempts per game. How does this compare to what the model predicts for this player?

Solution

- Circle the point at (2.1, 18.6).
- A point above the line represents a player who scores more points per game than predicted by their number of free throw attempts.
- 20.2 points per game, because $4.413(4.5) + 0.377$ is roughly equal to 20.2.
- The model predicts that with 4.1 free throw attempts per game, the player should score $4.413(4.1) + 0.377$, or about 18.5 points per game. That means the player is scoring less than the model predicts they should.

Lesson 5: Describing Trends in Scatter Plots

Goals

- Critique (orally and in writing) a given line of fit on a scatter plot, and draw a different linear model of the same data.
- 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—students won't study things like least-squares regression until high school) and begin to draw their own lines to fit data in a scatter plot (MP5).

Alignments

Addressing

- 8.SP.A.1: 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.
- 8.SP.A.2: 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.

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect
- MLR1: Stronger and Clearer Each Time
- MLR2: Collect and Display
- MLR7: Compare and Connect
- Notice and Wonder
- Think Pair Share

- Which One Doesn't Belong?

Required Materials

Dried linguine pasta

We specified linguine since it is flatter and less likely to roll around than spaghetti.

straightedge, but sometimes it is preferable to use an unruled straightedge, like a blank index card.

Straightedges

A rigid edge that can be used for drawing line segments. Sometimes a ruler is okay to use as a

Required Preparation

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

Student Learning Goals

Let's look for associations between variables.

5.1 Which One Doesn't Belong: Scatter Plots

Warm Up: 5 minutes

The purpose of this warm-up is to introduce students to positive and negative associations by comparing scatter plots with best-fit lines.

Addressing

- 8.SP.A.1
- 8.SP.A.2

Instructional Routines

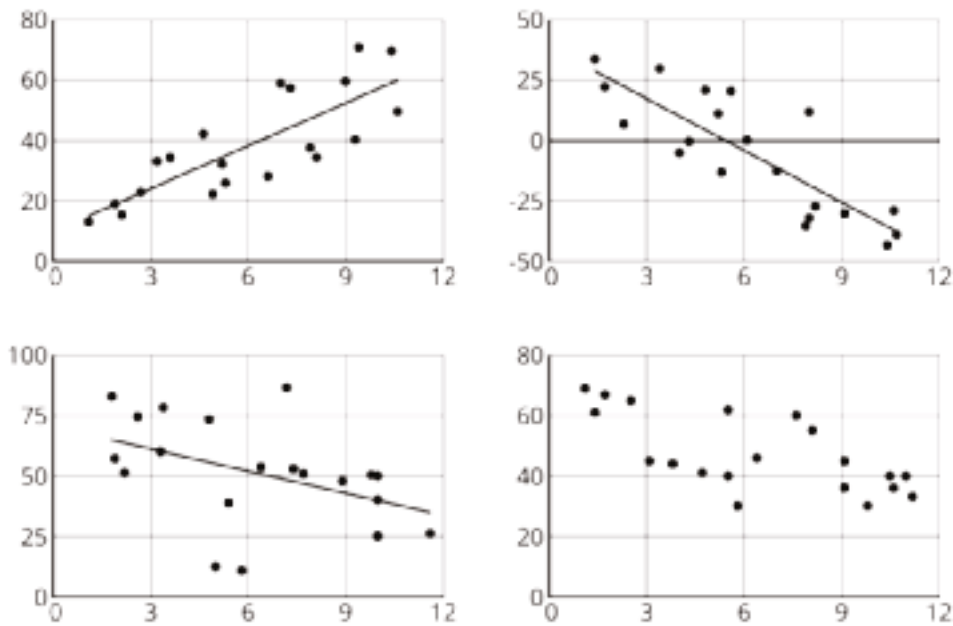
- Which One Doesn't Belong?

Launch

Display the image of the three scatter plots for all to see. Ask students to indicate when they have noticed one scatter plot that does not belong and can explain why it does not belong. Give students 1 minute of quiet think time.

Student Task Statement

Which one doesn't belong?



Student Response

Answers vary. Sample responses:

- The top left is the only one with a line that has a positive slope.
- The top right is the only one with negative y-coordinates.
- The bottom left has points that are much more spread out.
- The bottom right does not have a line to show a connection in the data.

Activity Synthesis

Poll the class for opinions on which scatter plot doesn't belong. Select one student for each scatter plot to explain their reasoning. Draw out reasons for each figure, attending to appropriate vocabulary and precise use of language (MP6). Record and display their reasoning for all to see, referencing the scatter plot when appropriate.

The bottom left plot can be contrasted using how well the line fits, but students do not have access to this language yet, so they might say something like, "The points are farther away from the line." The data points in the bottom left plot can also be described as "more variable" than the data in the other two scatter plots, which is a term that students might know from previous work in statistics.

During the discussion, introduce new vocabulary:

- A scatter plot that can be modeled by a line with a positive slope can be said to show a *positive trend* or a **positive association**.

- A scatter plot that can be modeled by a line with a negative slope can be said to show a *negative trend* or a **negative association**.

5.2 Fitting Lines

20 minutes (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).

While students are working, monitor for different approaches to deciding where to draw the line. Approaches might include:

- Drawing the line that connects the leftmost point to the rightmost point (or the topmost point to the bottommost point).
- Drawing the line that passes directly through as many points as possible.
- Making a visual estimate of an appropriate slope, and then ensuring half the points lie above the line and half the points lie below the line.

Addressing

- 8.SP.A.2

Instructional Routines

- Anticipate, Monitor, Select, Sequence, Connect
- MLR7: Compare and Connect

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.

If using the digital activity, students can work in groups of two and create their lines of best fit digitally.

Access for Students with Disabilities

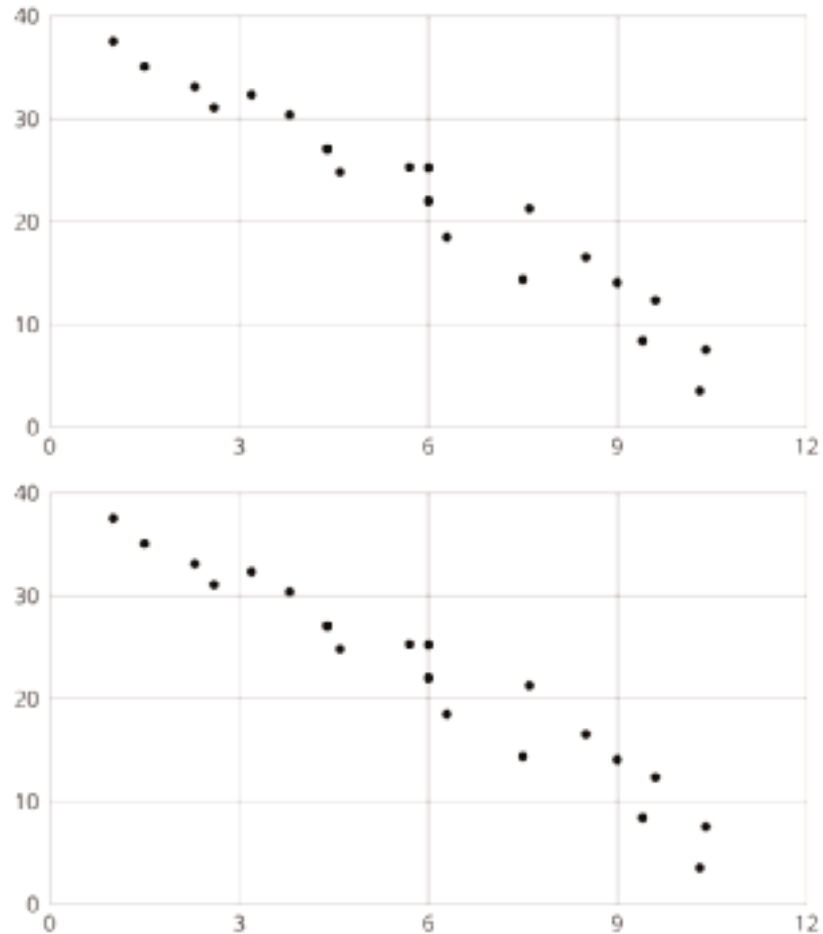
Action and Expression: Develop Expression and Communication. To help get students started sharing their ideas and describing what makes a line fit a data set well, display sentence frames such as “We are trying to...,” “Let’s try...,” and “I noticed ____ so I...”

Supports accessibility for: Language; Organization

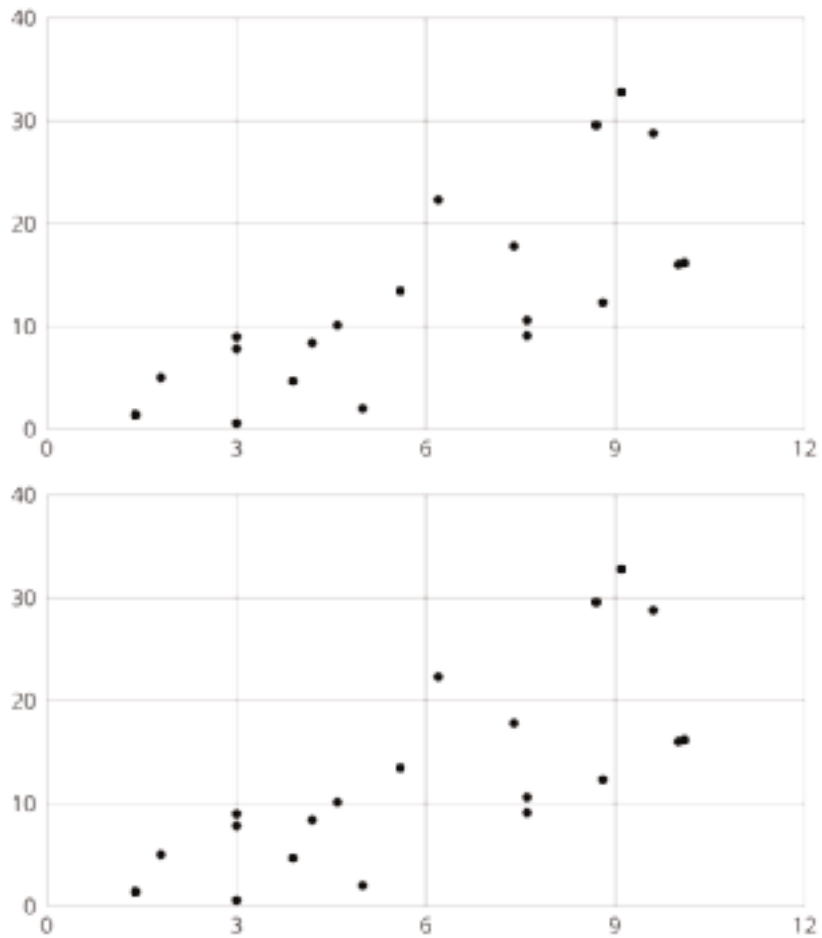
Student Task Statement

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. Pick the line that you think best fits the data. Compare it with a partner's.



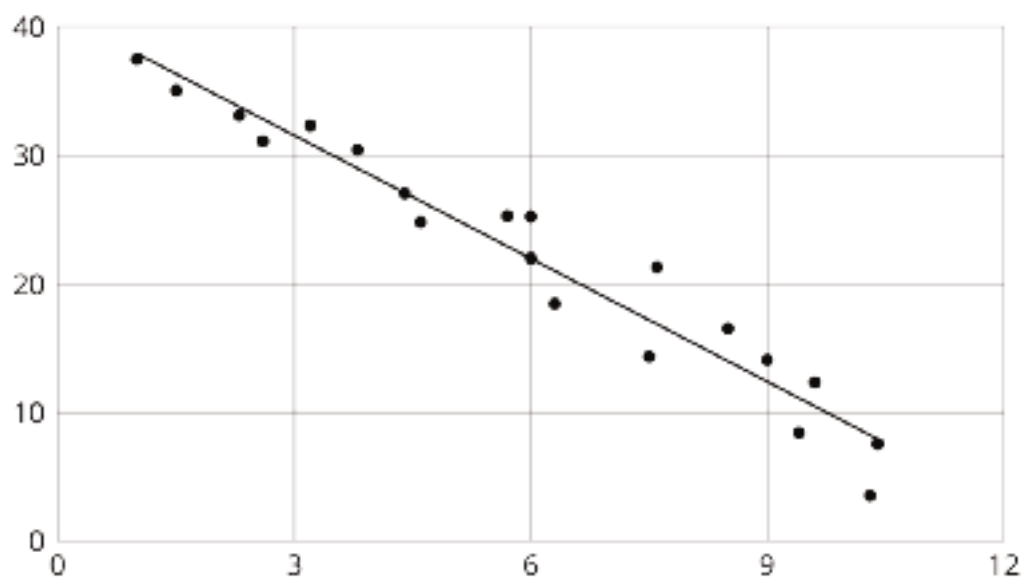
2. Here are two copies of another scatter plot. Experiment with drawing lines to fit the data. Pick the line that you think best fits the data. Compare it with a partner's.



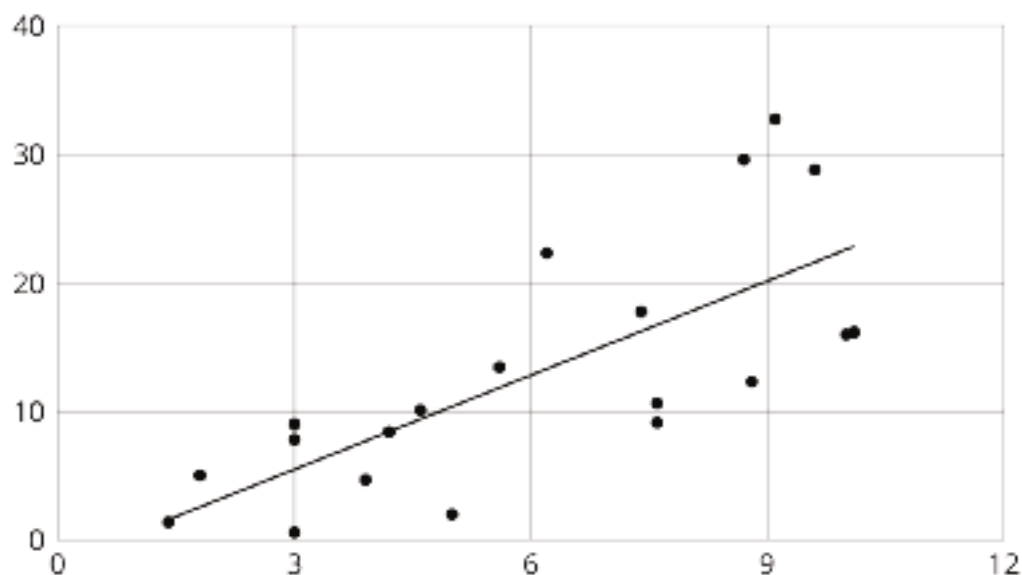
3. In your own words, describe what makes a line fit a data set well.

Student Response

1. Answers vary. 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. Answers vary. 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. Answers vary. Possible responses:
- 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

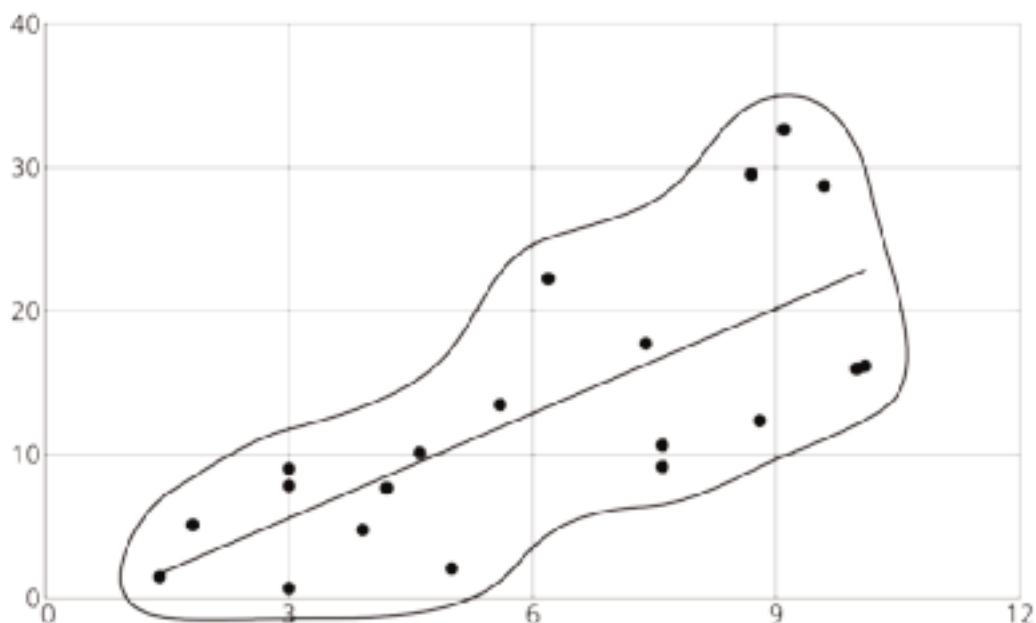
The purpose of this discussion is to look at some strategies for drawing a line that fits the data well.

Select students to share their approaches to drawing a line to fit the data using the sequence listed in the Activity Narrative. If any of these strategies are not represented in student's work, bring them up as possibilities for the class to critique.

The first two approaches often result in a model that does not fit the data well. Display these examples and select students to connect how the lines do and do not fit the data. Focus on the last approach as a preferred method. Emphasize that when deciding how to draw a line to model a scatter plot, it's important to consider the whole data set, not just a few points.

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.



Access for English Language Learners

Speaking, Listening: MLR7 Compare and Connect. Use this routine when students present their strategies for drawing lines that fit the data well. Ask students to first identify “what is the same and what is different” about each approach. Draw students’ attention to the different ways the lines were constructed (e.g., connecting leftmost and rightmost points; drawing a line through as many points as possible; drawing a line where half the points fall above and below the line, etc.). In this discussion, emphasize the mathematical language used to make sense of the different ways to construct lines that fit the data well. These exchanges strengthen students’ mathematical language use and reasoning when constructing and analyzing lines that fit data points well.

Design Principle(s): Maximize meta-awareness

5.3 Good Fit Bad Fit

Optional: 15 minutes

If students understand what makes a 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.

Addressing

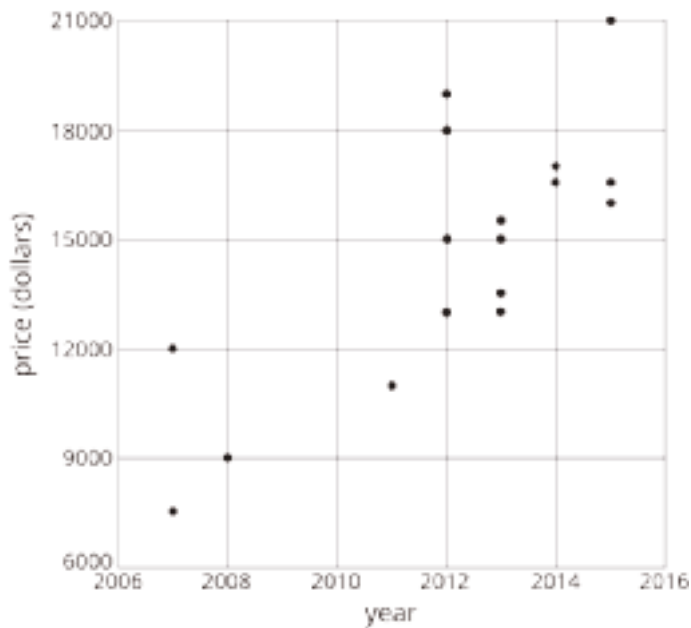
- 8.SP.A.2

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?”



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?

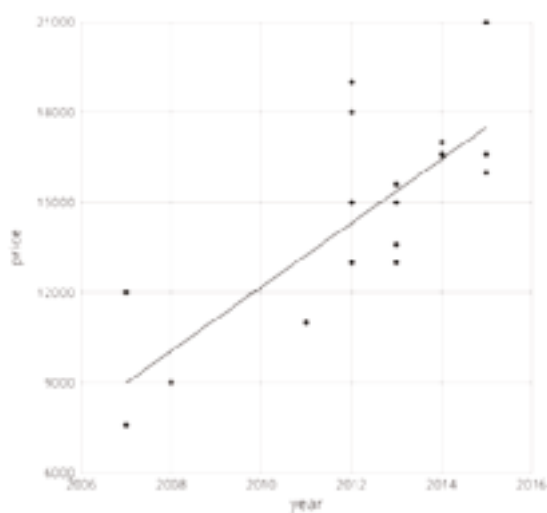
Tell students that these are all prices of used cars that are all the same make and model that are for sale. For each car, the scatter plot shows its year of manufacture and the price at which it is being sold. Ask students a few questions to familiarize themselves with the graph, like:

- "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.)

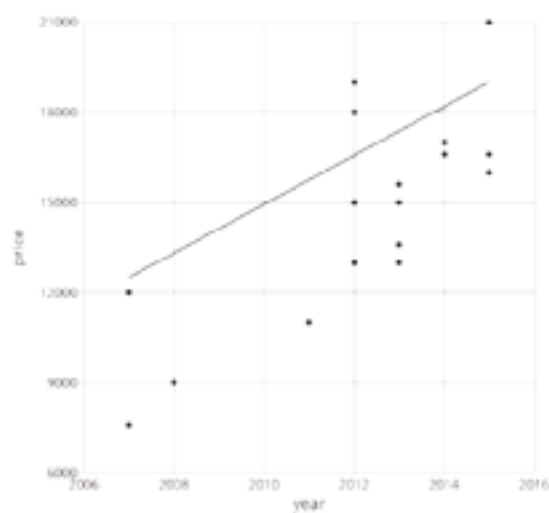
Tell students that in this task, they are going to see two different models for this set of data.

Student Task Statement

The scatter plots both show the year and price for the same 17 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

1. Answers vary. Sample response: Four points are fairly close to the line, so 4 prices are well predicted. The model underestimates the price of 5 cars and overestimates the price of 8 cars.
2. Answers vary. Sample response: No points are very close to the line, so none is very well predicted. The model underestimates the price of 3 cars and overestimates the price of 14 cars.

3. Answers vary. Sample response: Model A over or underpredicts the price of 3 cars by about \$3,000 or more. For model B, there are 5 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. 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

Writing, Speaking: MLR1 Stronger and Clearer Each Time. Use this routine to give students a structured opportunity to revise and refine 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 them strengthen their ideas and clarify their language (e.g., "Why do you think model A (or B) is better?", "How did you compare the two models?", and "Can you give an example?", etc.). Students can borrow ideas and language from each partner to strengthen their final version.

Design Principle(s): Optimize output (for explanation)

5.4 Practice Fitting Lines

10 minutes

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.

Addressing

- 8.SP.A.2

Instructional Routines

- MLR2: Collect and Display
- Think Pair Share

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.

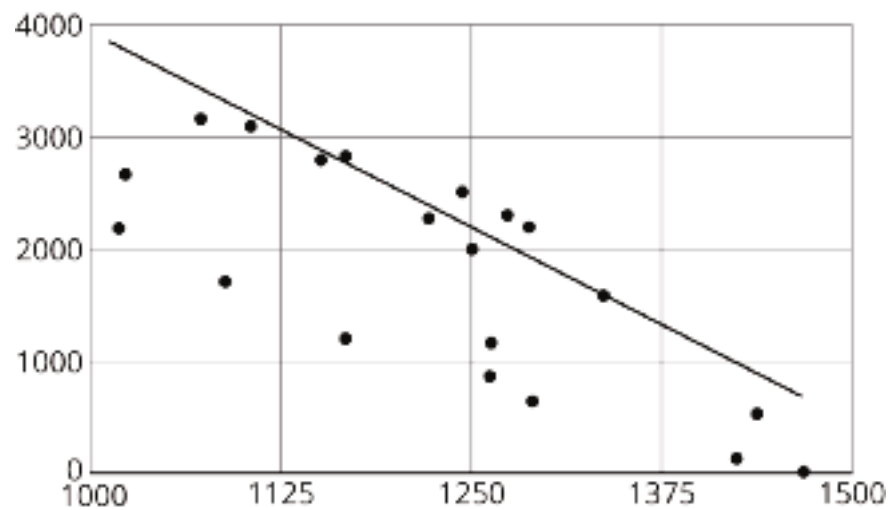
Access for English Language Learners

Conversing, Representing, Writing: MLR2 Collect and Display. As students, listen for and collect the vocabulary, gestures, and phrases students use to critique and improve a line that fits the data well. Write the students' words and gestures on a visual display that can be referenced in future discussions. Continue to update collected student language throughout the entire activity. Remind students to borrow language from the display as needed.

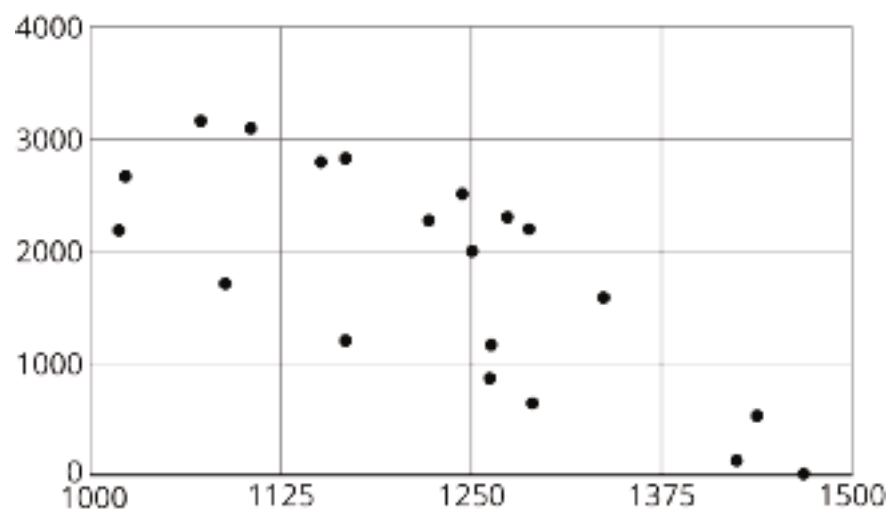
Design Principle(s): Optimize output (for justification); Maximize meta-awareness

Student Task Statement

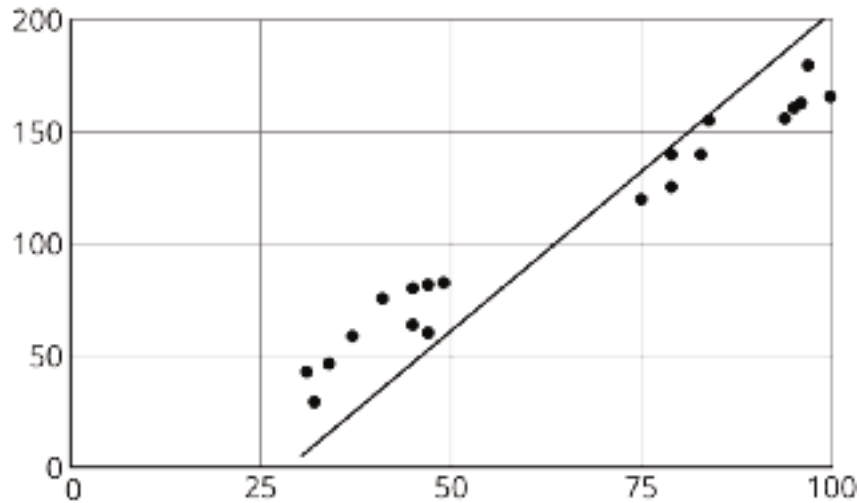
1. Is this line a good fit for the data? Explain your reasoning.



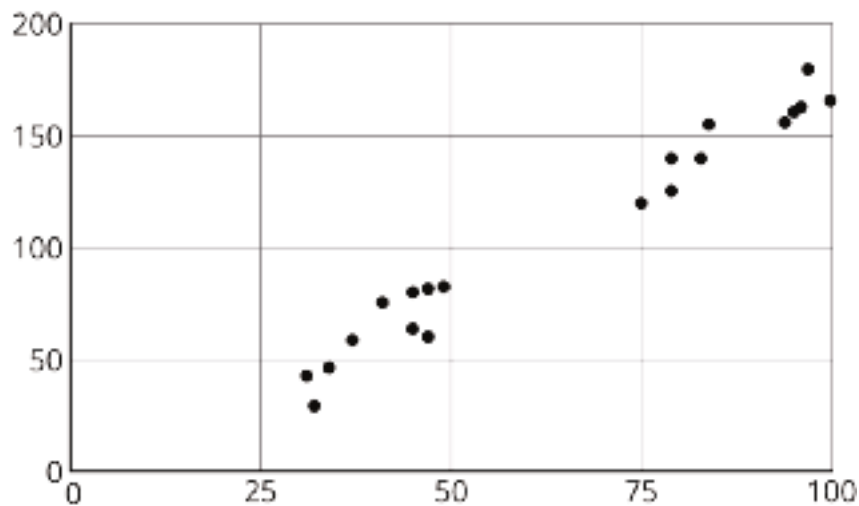
2. Draw a line that fits the data better.



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

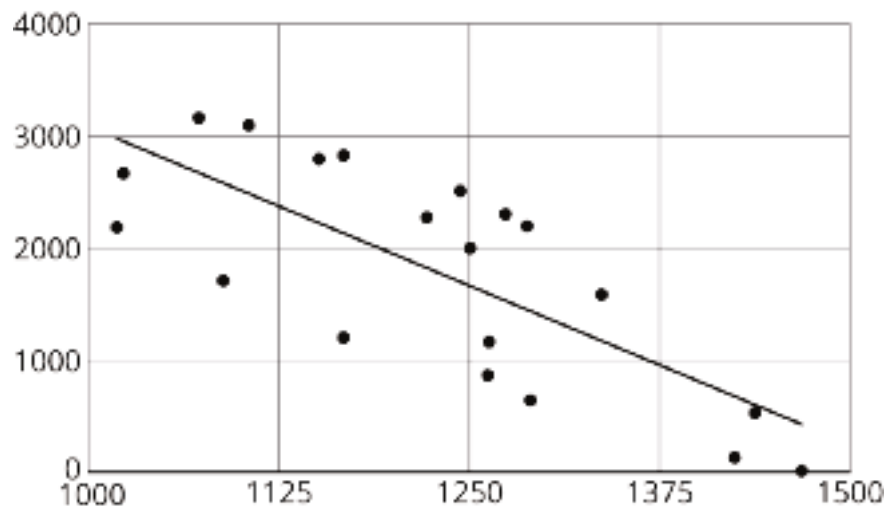


4. Draw a line that fits the data better.

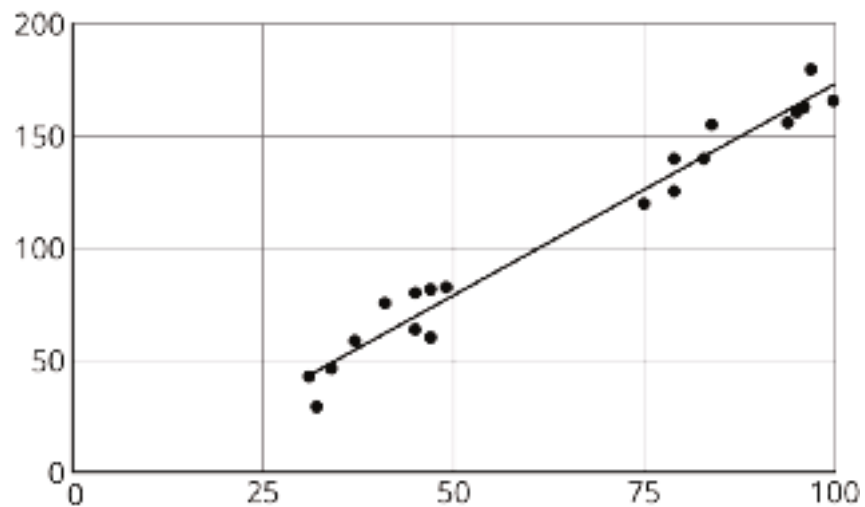


Student Response

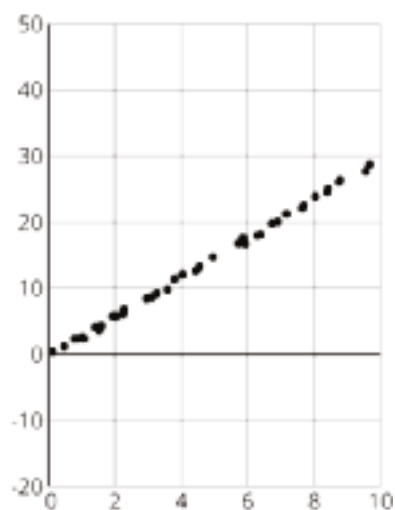
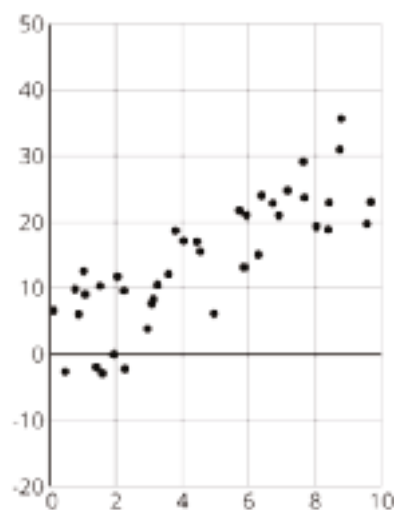
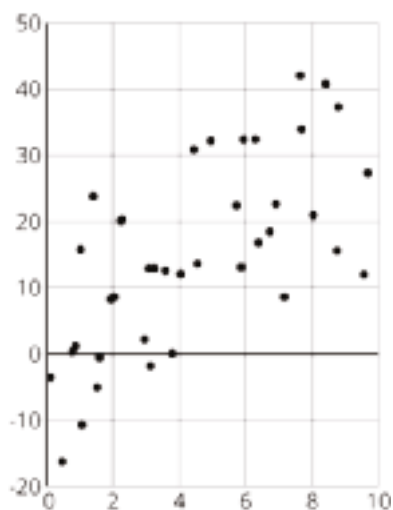
1. No. There are many more points below the line than above it. A line with a good fit should move down the graph from where this line is drawn.
2. Answers vary. An appropriate line created by students will have a negative trend and go "through the middle" of the cloud of points. For reference, here is a line made using a least-squares regression:



3. Not great. The line could be tilted to have a slightly lesser slope to fit the line better.
4. Answers vary. 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:



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 -2 and 2 to the y -coordinate. The third scatter plot added a random number between -10 and 10 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.

Student Response

Answers vary. Sample response:

1. The line $y = 3x$ or similar would be a good fit for each graph.

2. Since the points were closer together in the first two graphs, it was easier to draw the line for those two graphs. 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

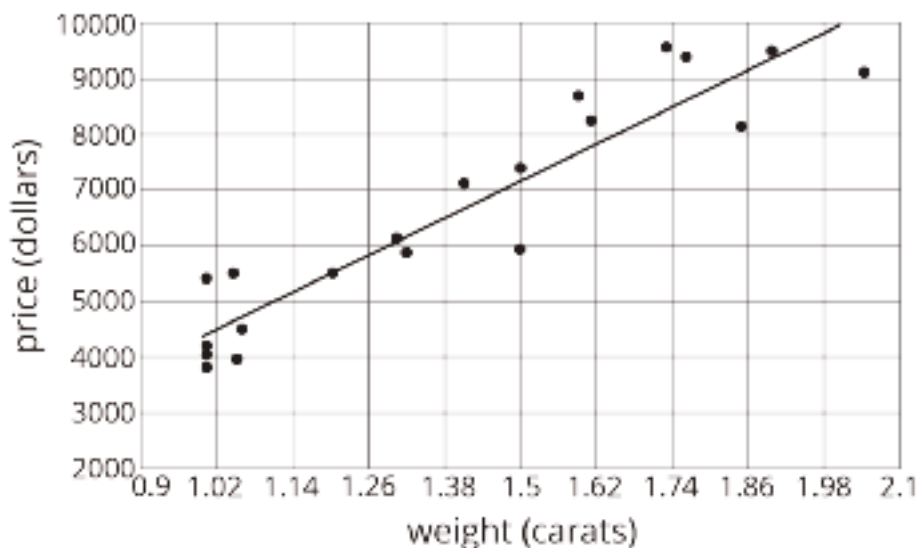
The purpose of the discussion is for students to recognize the important aspects of a linear model for a set of data.

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 weight of a diamond and its price?" (There is a positive association. Or, as the weight increases, the price 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. The points are as close as possible to the line.)

5.5 This is One Way to Do It

Cool Down: 5 minutes

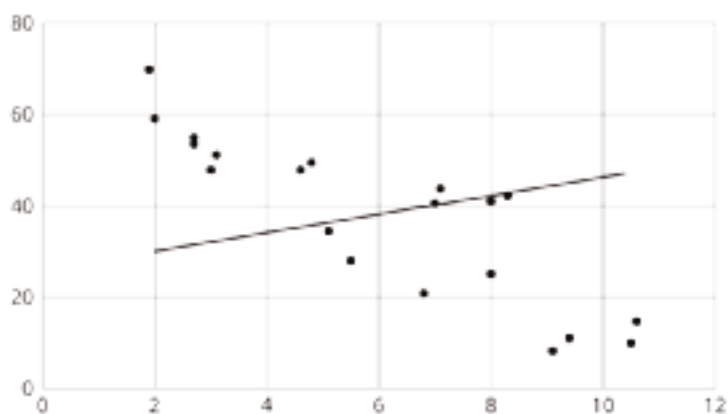
This activity addresses some misconceptions students may have while thinking about linear models. Students critique attempts to draw a line that fits data in a scatter plot. In particular, two lines are drawn that resolve some concerns when fitting lines but disregard others.

Addressing

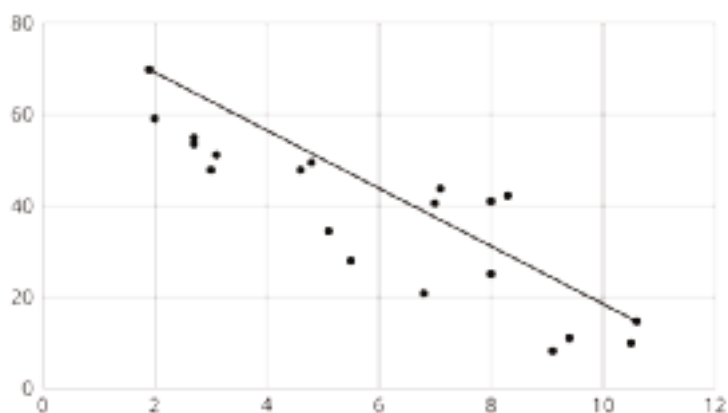
- 8.SP.A.2

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.



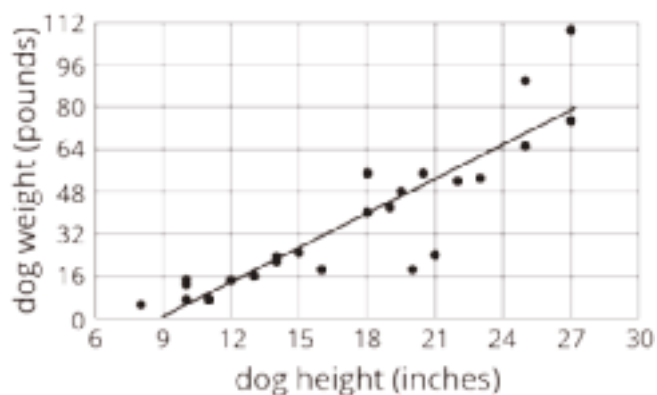
Student Response

1. Disagree. Explanations vary. 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. Explanations vary. Sample responses: The line is not a good fit because most of the points are below it. The line is not a good fit because the trend of the scatter plot is steeper than the slope of the graph.

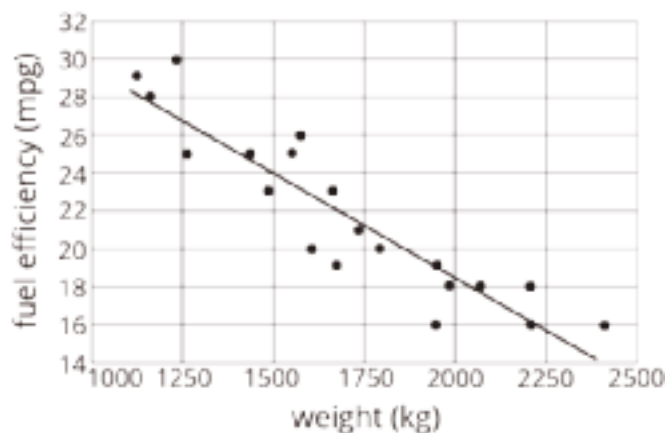
Student Lesson Summary

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.



Because the model fits the data well and because the slope of the line is positive, we say that there is a **positive association** between dog height and dog weight.

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



Because the slope of a line that fits the data well is negative, we say that there is a **negative association** between the fuel efficiency and weight of a car.

Glossary

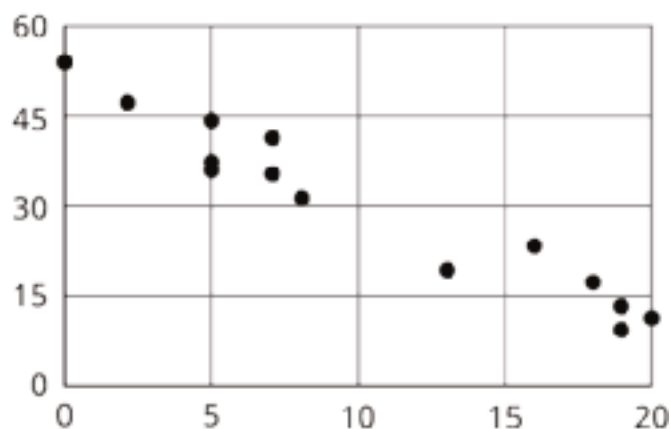
- negative association
- positive association

Lesson 5 Practice Problems

Problem 1

Statement

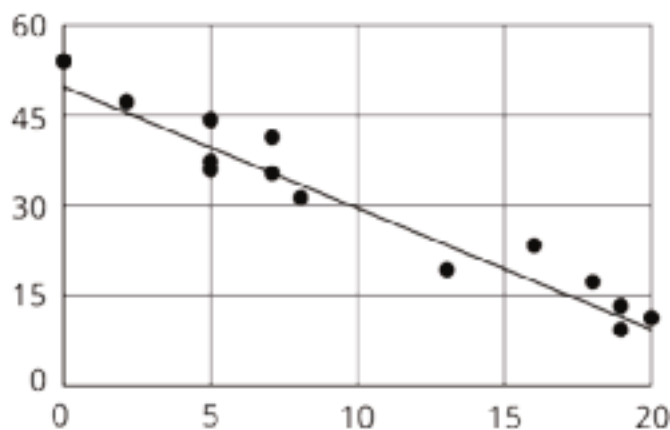
- a. Draw a line that you think is a good fit for this data. For this data, the inputs are the horizontal values, and the outputs are the vertical values.



- b. Use your line of fit to estimate what you would expect the output value to be when the input is 10.

Solution

- a. Answers vary. Sample response:

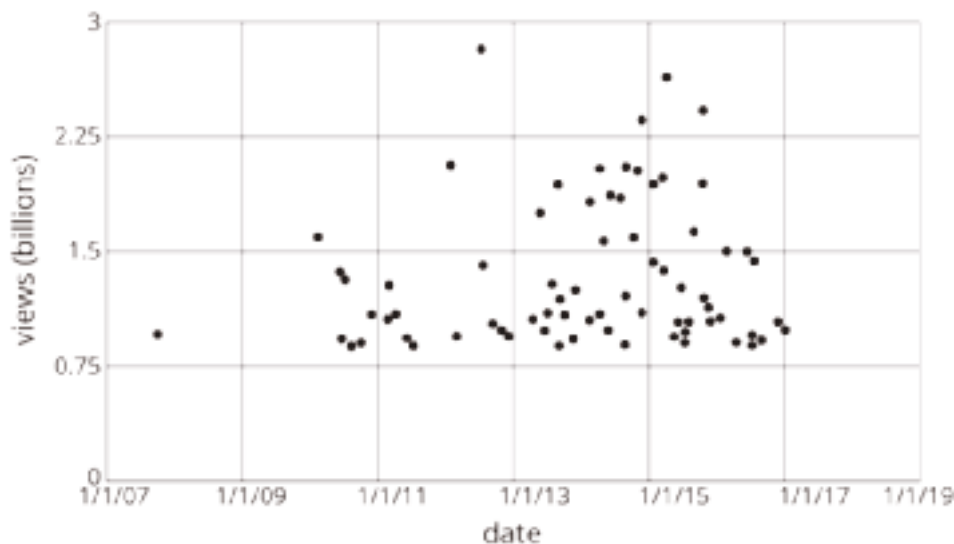


- b. Answers vary. Sample response: The output would be close to 30.

Problem 2

Statement

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



- Use the scatter plot to estimate the number of views for the most popular video in this 10-year span.
- Estimate when the 4th most popular video was released.

Solution

- The most popular video has roughly 2.8 billion views.
- Late 2014

(From Unit 6, Lesson 3.)

Problem 3

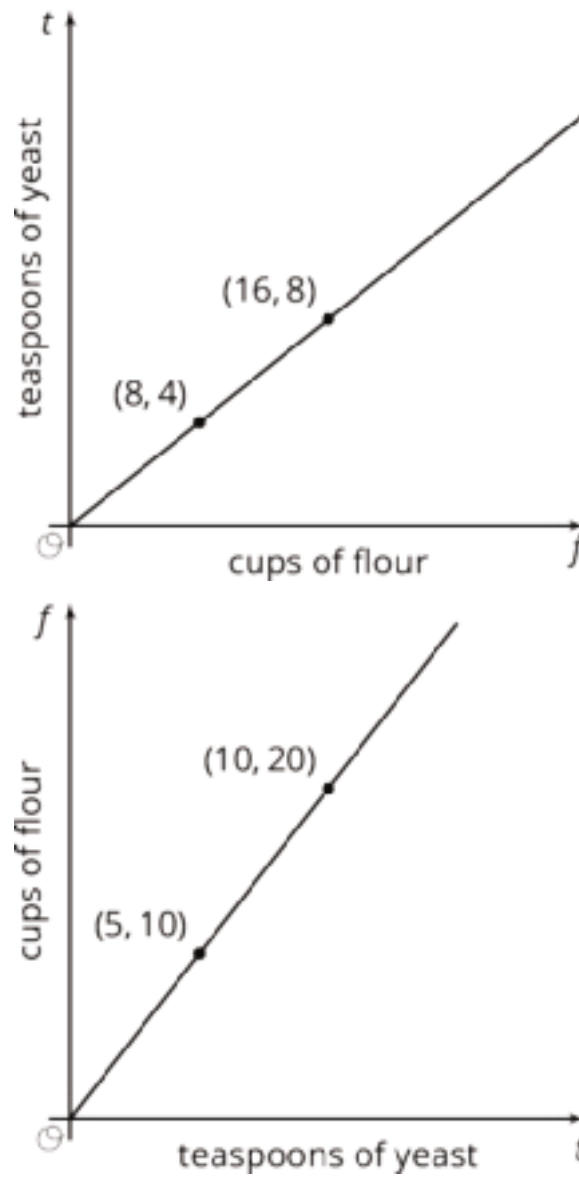
Statement

A recipe for bread calls for 1 teaspoon of yeast for every 2 cups of flour.

- Name two quantities in this situation that are in a functional relationship.
- Write an equation that represents the function.
- Draw the graph of the function. Label at least two points with input-output pairs.

Solution

- The amount of yeast and the amount of flour are in a functional relationship.
- Let t represent the number of teaspoons of yeast and f represent the number of cups of flour. If the amount of flour is treated as a function of the amount of yeast, then the equation is $f = 2t$. If it's the other way around, the equation is $t = \frac{1}{2}f$.
- Points plotted vary. Sample responses:



(From Unit 5, Lesson 8.)

Lesson 6: The Slope 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.
- 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 slope of a line fit to data in context.

Learning Targets

- I can use the slope 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 (MP6). Students also interpret the slopes of fitted lines in context (MP2) and identify positive and negative associations of scatter plots without a fitted line shown (MP7).

Alignments

Building On

- 8.EE.B.6: Use similar triangles to explain why the slope 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 $y = mx + b$ for a line intercepting the vertical axis at b .

Addressing

- 8.SP.A.1: 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.
- 8.SP.A.2: 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.
- 8.SP.A.3: Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope 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

- 8.SP.A.3: Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope 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

- MLR1: Stronger and Clearer Each Time
- MLR8: Discussion Supports
- Think Pair Share

Student Learning Goals

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

6.1 Estimating Slope

Warm Up: 5 minutes

The purpose of this warm-up is for students to estimate the slope 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.

Building On

- 8.EE.B.6

Building Towards

- 8.SP.A.3

Instructional Routines

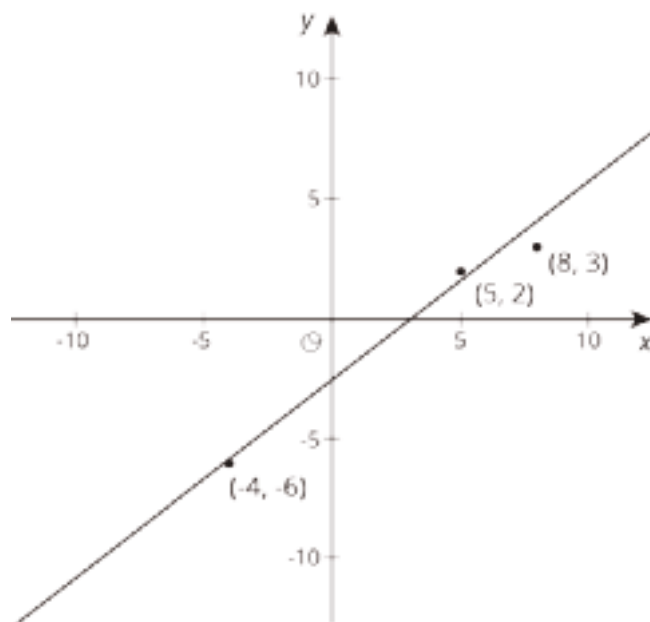
- Think Pair Share

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 slope of the line.



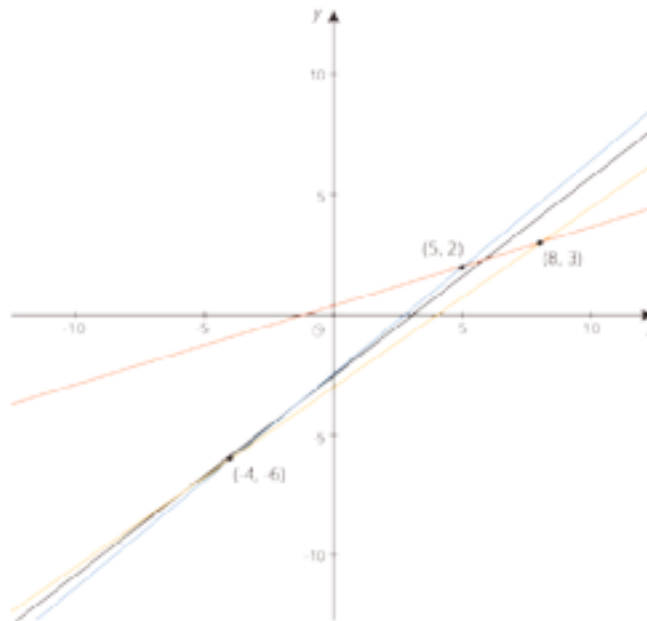
Student Response

The slope is a little less than $\frac{8}{9}$, which is the slope of the segment connecting $(-4, -6)$ with $(5, 2)$ since $\frac{2-(-6)}{5-(-4)} = \frac{8}{9}$. The slope is a little more than $\frac{9}{12} = \frac{3}{4}$, which is the slope of the 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 slope 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 slope, consider displaying this graph for all to see:



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

6.2 Describing Linear Associations

10 minutes

Students practice using precise wording (MP6) to describe the positive or negative association between two variables given scatter plots of data.

Addressing

- 8.SP.A.1
- 8.SP.A.2

Building Towards

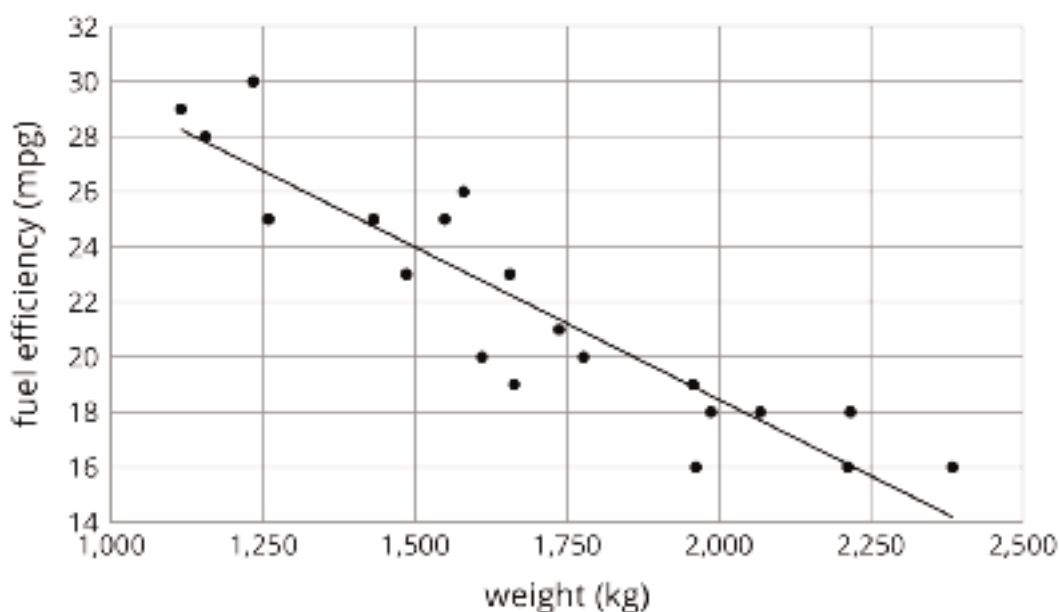
- 8.SP.A.3

Instructional Routines

- MLR8: 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.

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.”

Access for English Language Learners

Speaking: MLR8 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; Optimize output (for explanation)

6.3 Interpreting Slopes

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 slope. The numerical value of the slope is then interpreted in the context of the problem (MP2).

Addressing

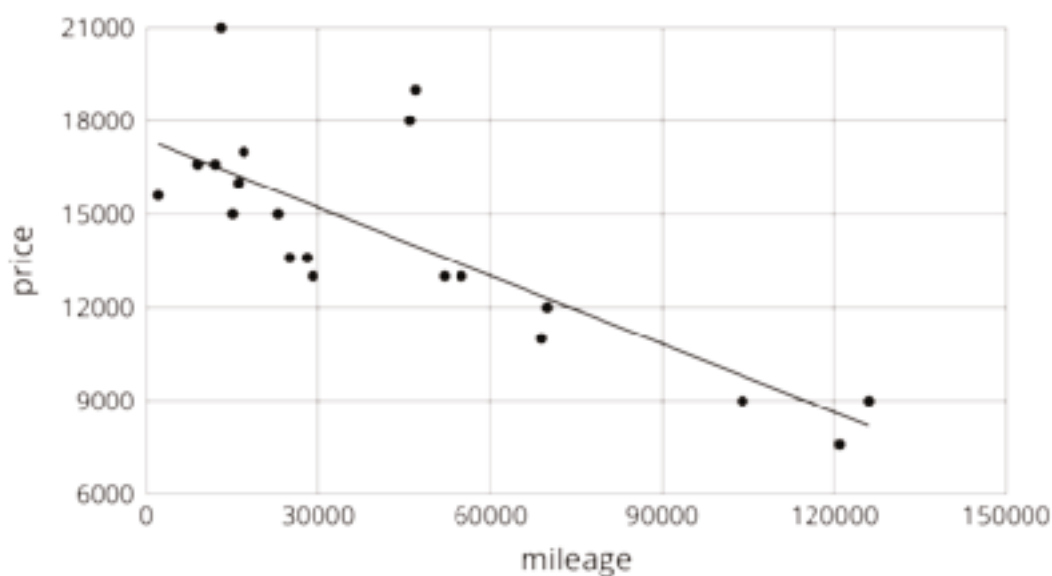
- 8.SP.A.3

Instructional Routines

- MLR8: 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 slope of the line as -0.073 . Ask students to think about what that slope 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 slope of the model means for three different sets of data.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. Chunk this task into more manageable parts to support students who benefit from support with organization and problem solving. For example, present only one scatter plot at a time.

Supports accessibility for: Organization; Attention

Access for English Language Learners

Speaking, Writing: MLR8 Discussion Supports. Provide a sentence frame to scaffold students' language when describing the meaning of slope in each representation. For example, "If the _____ increases by 1 _____, the model predicts that _____ (increases/decreases) by _____."

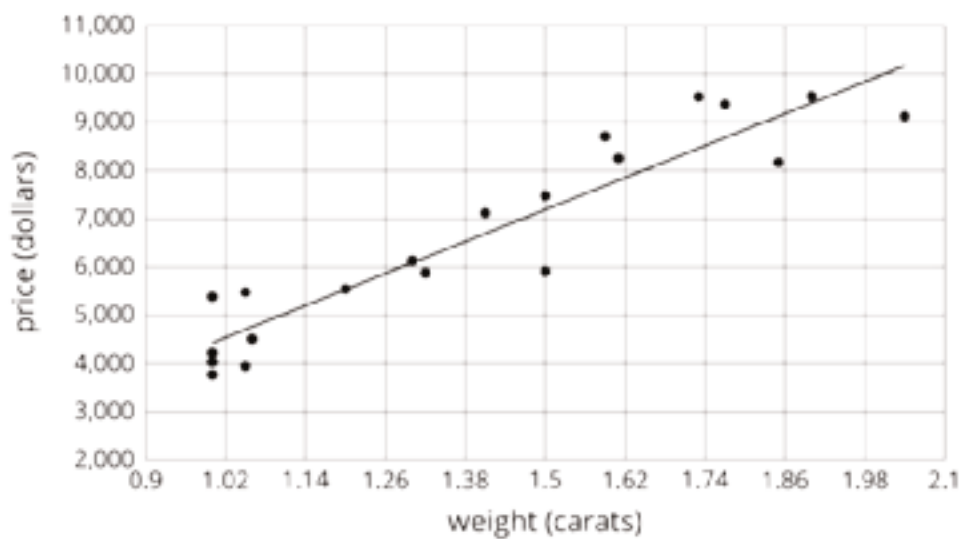
Design Principle: Maximize meta-awareness; Optimize output (for explanation)

Student Task Statement

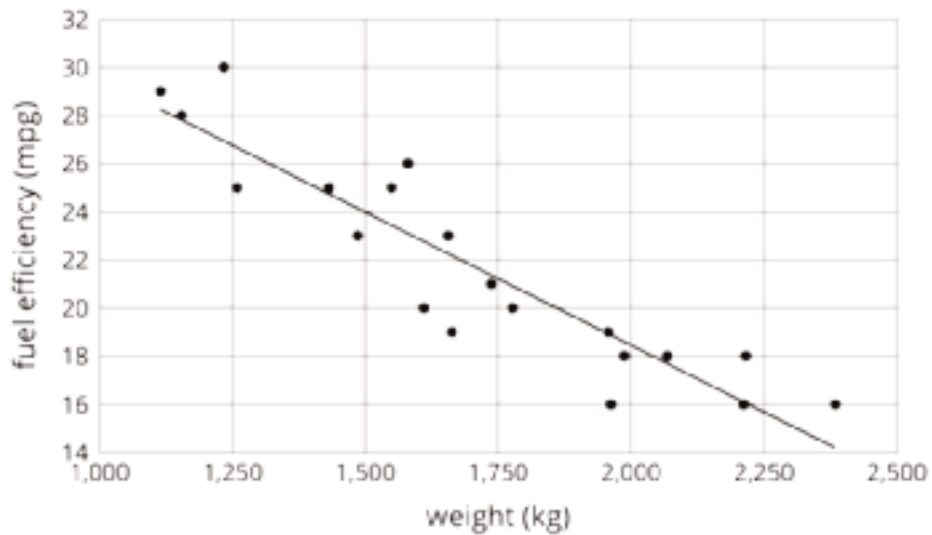
For each of the situations, a linear model for some data is shown.

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

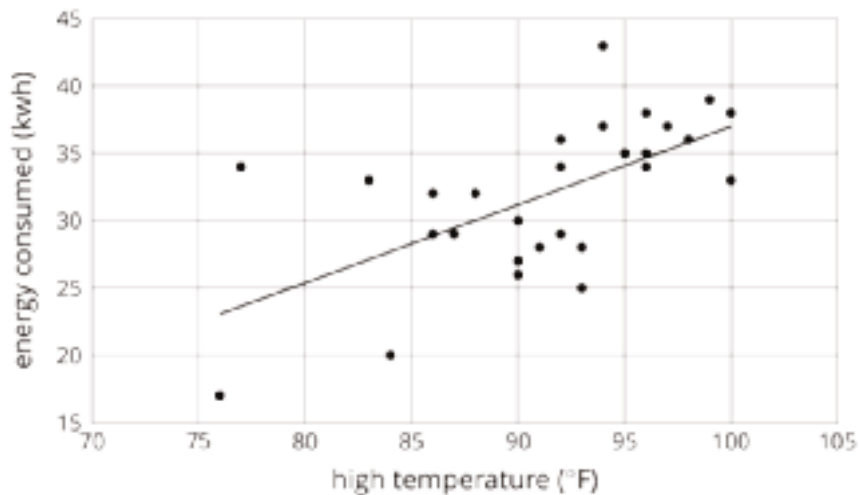
$$y = 5,520.619x - 1,091.393$$



$$y = -0.011x + 40.604$$



$$y = 0.59x - 21.912$$



Student Response

For diamond price vs size:

1. The slope 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:

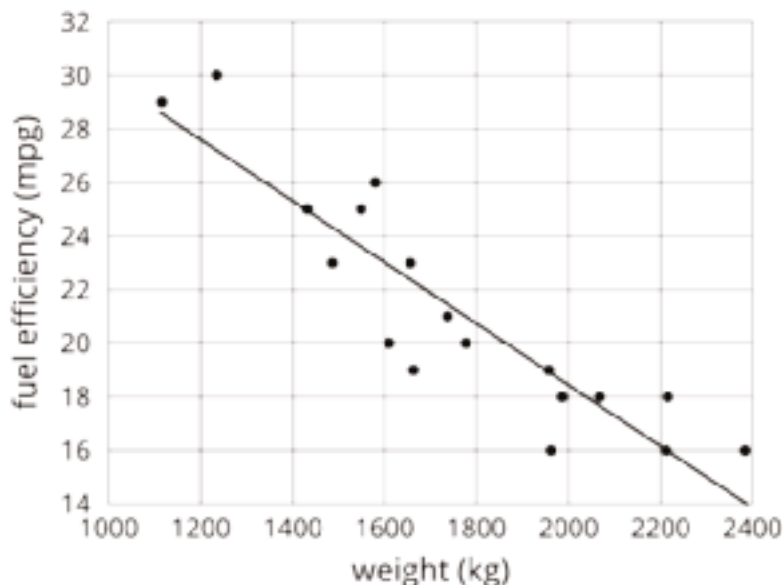
1. The slope is -0.011 .
2. If the weight of a car increases by 1 kg, the model predicts that its fuel efficiency decreases by 0.011 mpg.

Energy consumed vs daily high temperature:

1. The slope 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 $y = -0.0114x + 41.3021$.



1. What is the value of the slope 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 slope. 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 $-0.0114 \cdot 100 + 41.3021 \approx 40.16$.

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 slope for each situation?" (The coefficient of the x -coordinate in the equation.)
- "Do the answers for the meaning of the slopes 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 slope and negative slope in your interpretations?" (A positive slope means both variables increase together. A negative slope 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 (MP4).

Addressing

- 8.SP.A.2

Instructional Routines

- MLR1: 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 slope a linear model of the data would have.

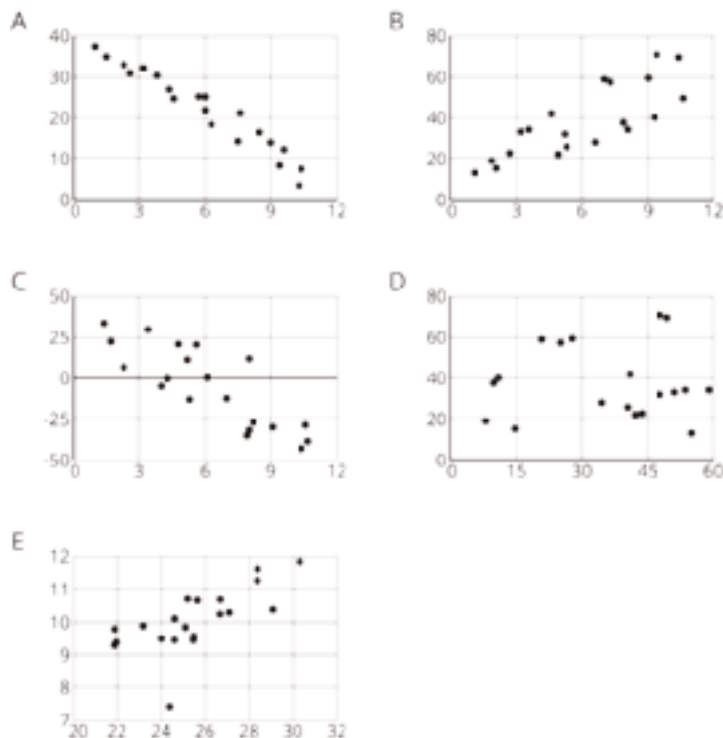
Access for Students with Disabilities

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 slope, a negative slope, or a slope of zero?



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 slope.
B: A linear model would fit and have a positive slope.
C: A linear model would fit and have a negative slope.
D: A linear model would not fit very well.
E: A linear model would fit and have a positive slope, 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 slope 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?”

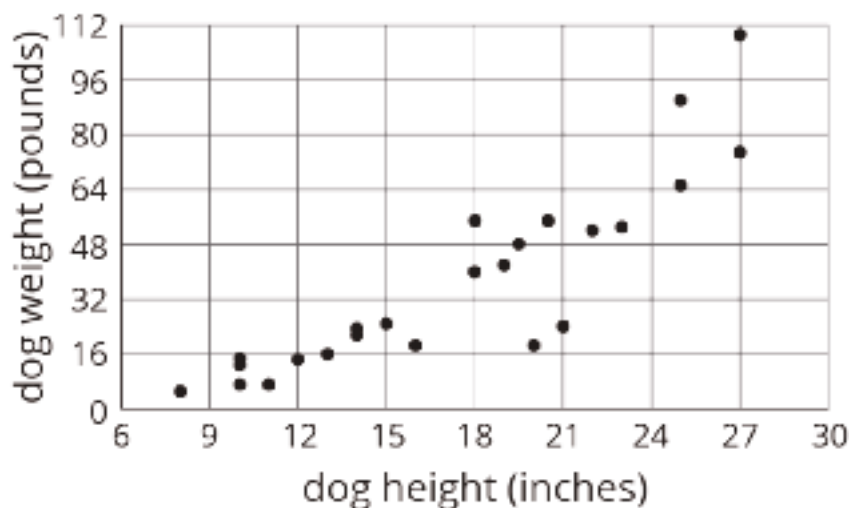
Access for English Language Learners

Writing: MLR1 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): Optimize 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 slope 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 slope (negative slope, respectively).)
- “The slope 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

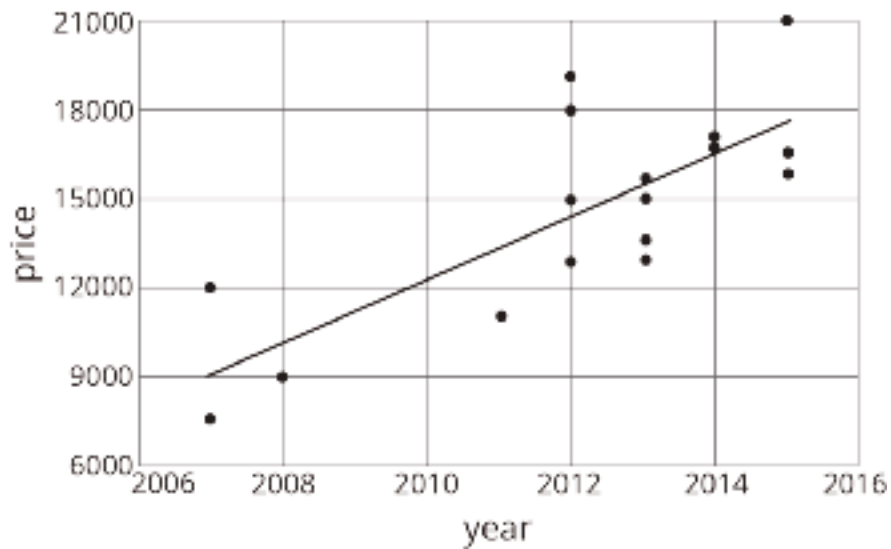
Students use a scatter plot of data representing used car prices and their model year to determine the sign of the slope of a linear model. Then students interpret this slope in the context of the data.

Addressing

- 8.SP.A.3

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.



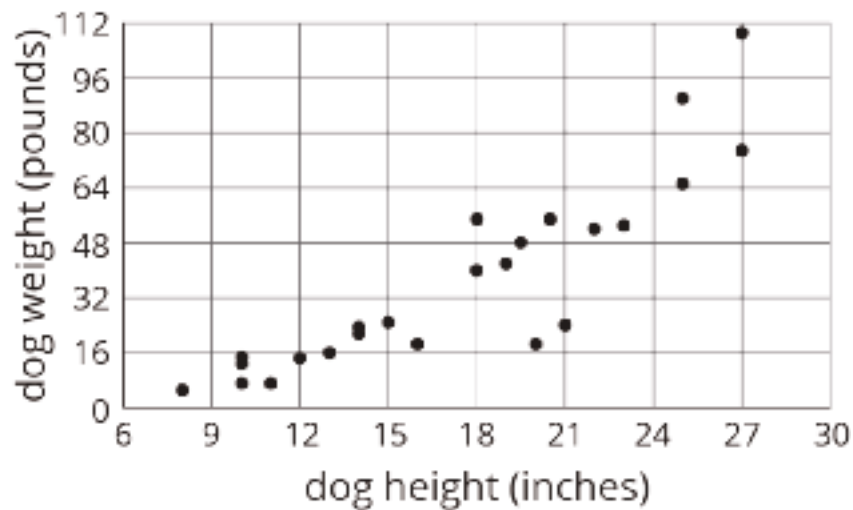
1. Is the slope positive or negative?
2. Which of these values is closest to the slope of the linear model shown in the scatter plot?
 - 1,000
 - 3,000
 - -1,000
 - -3,000
3. Use the value you selected to describe the meaning of the slope in this context.

Student Response

1. The slope 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

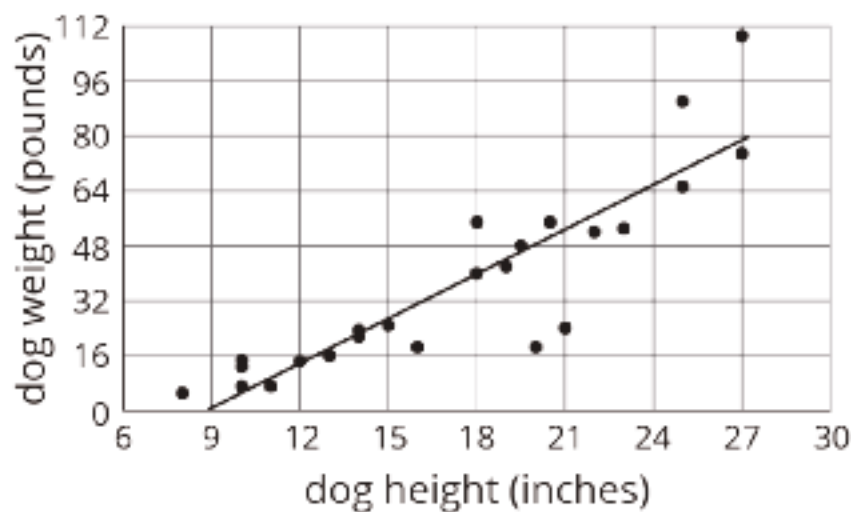
Here is a scatter plot that we have seen before. As noted earlier, we can see from the scatter plot that taller dogs tend to weigh more than shorter dogs. Another way to say it is that weight tends to increase as height increases. When we have a positive association between two variables, an increase 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 slope. For example, the equation of the fitted line is

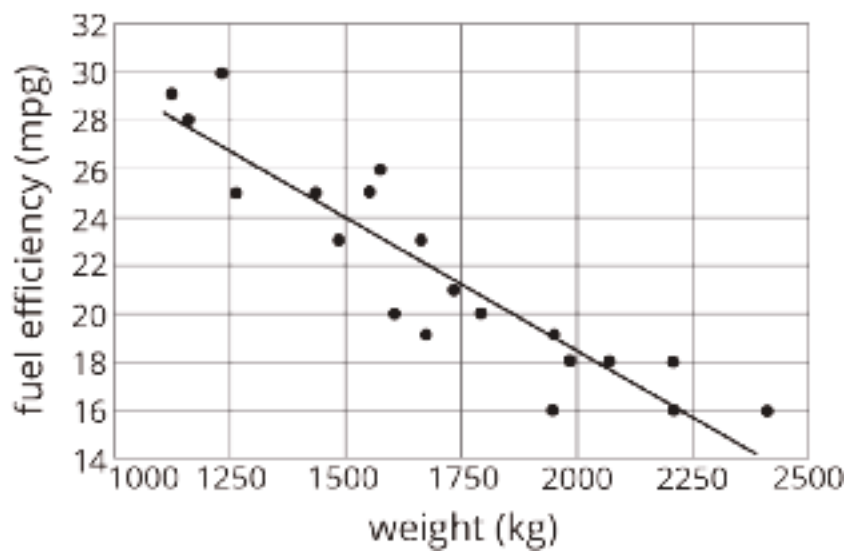
$$w = 4.27h - 37$$

where h is the height of the dog and w is the predicted weight of the dog.



The slope 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 slope of the fitted line shown is -0.01.



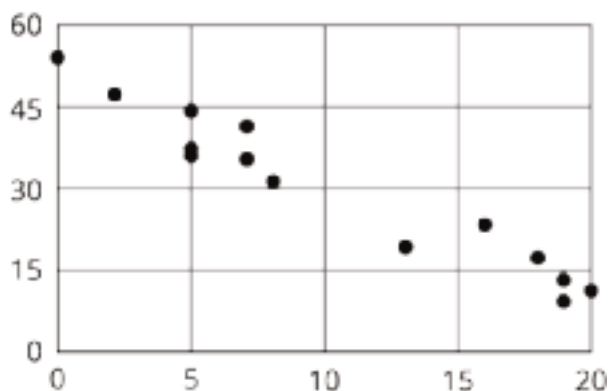
This tells us that for every 1-kilogram increase in the weight of the car, the fuel efficiency is predicted to decrease by 0.01 miles per gallon. When we have a negative association between two variables, an increase in one means there tends to be a decrease in the other.

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.

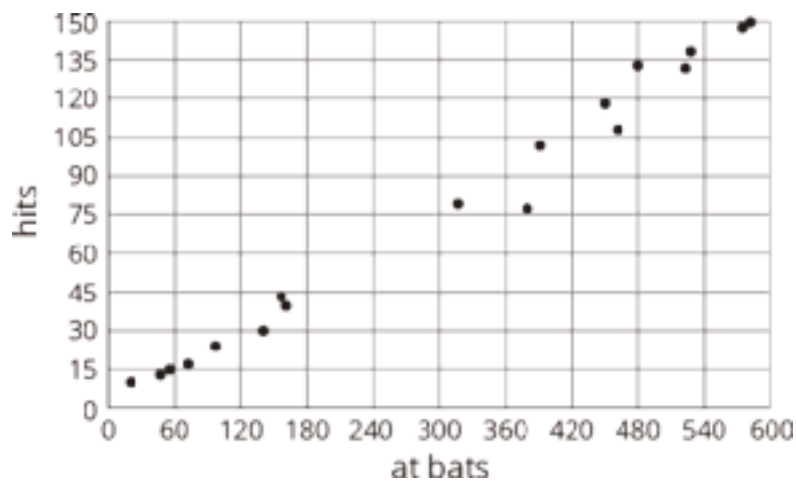
Solution

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.

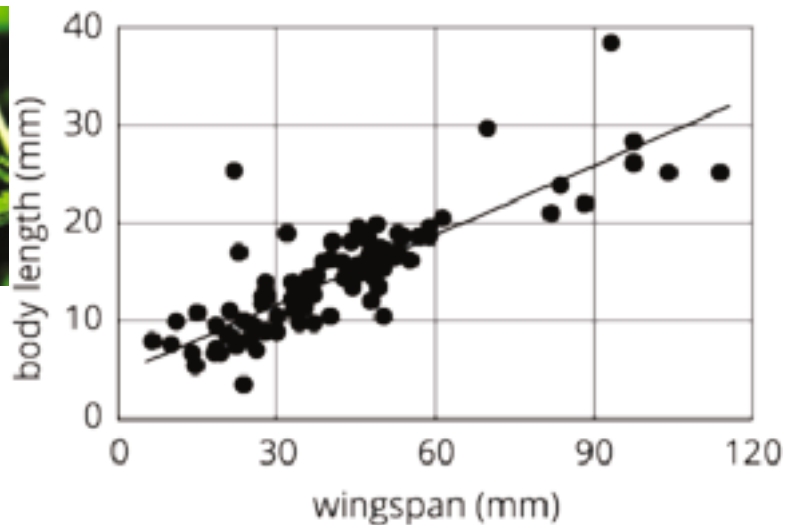
Solution

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 slope 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.

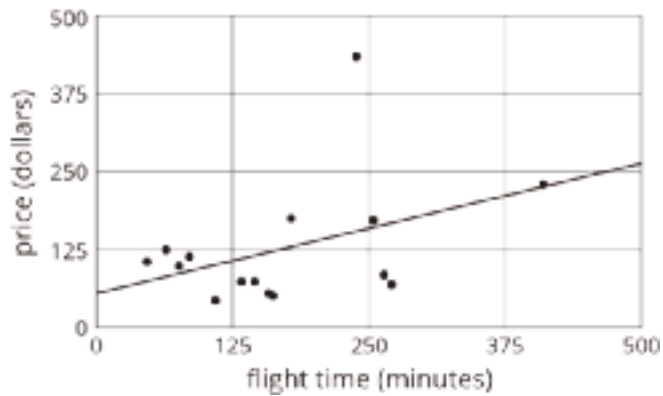
Solution

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.



- Circle any data that appear to be outliers.
- Use the graph to estimate the difference between any outliers and their predicted values.

Solution

- The point at (239, 436) appears to be an outlier.
- This point represents a destination that costs around \$250 more than the model predicts for its flight time.

(From Unit 6, Lesson 4.)

Problem 5

Statement

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

Solution

(12, -23)

(From Unit 4, Lesson 14.)

Lesson 7: Observing More Patterns in Scatter Plots

Goals

- Categorize data sets, and describe (orally) the properties used to create categories.
- Describe (orally) features of data on scatter plots, including “linear” and “nonlinear association” and “clustering” using informal language.
- Explain (orally) what might cause a nonlinear association or clustering of data points in context.

Learning Targets

- I can pick out clusters in data from a scatter plot.
- I can use a scatter plot to decide if two variables have a linear association.

Lesson Narrative

In this lesson, students see non-linear associations for the first time and do a card sort to distinguish linear and non-linear associations. They visually identify clusters in data. Students will not study non-linear associations or clustering using quantitative tools; instead they will rely on visual patterns in scatter plots (MP7).

Alignments

Addressing

- 8.SP.A.1: 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.

Instructional Routines

- MLR2: Collect and Display
- MLR8: Discussion Supports
- Notice and Wonder
- Take Turns
- Think Pair Share

Required Materials

Pre-printed cards, cut from copies of the Instructional master

Required Preparation

Print and cut up cards from the Scatter Plot City Instructional master. Prepare 1 set of cards for every student.

Student Learning Goals

Let's look for other patterns in data.

7.1 Notice and Wonder: Nonlinear Scatter Plot

Warm Up: 5 minutes

The purpose of this warm-up is for students to analyze a scatter plot by noticing things and asking questions (MP1). This scatter plot gives students the opportunity to observe a relationship that is not described well with a linear model. As production increases, price tends to decrease.

However, there is a lower limit to the price, so the model tends to level out after a certain amount of production is reached.

Addressing

- 8.SP.A.1

Instructional Routines

- Notice and Wonder

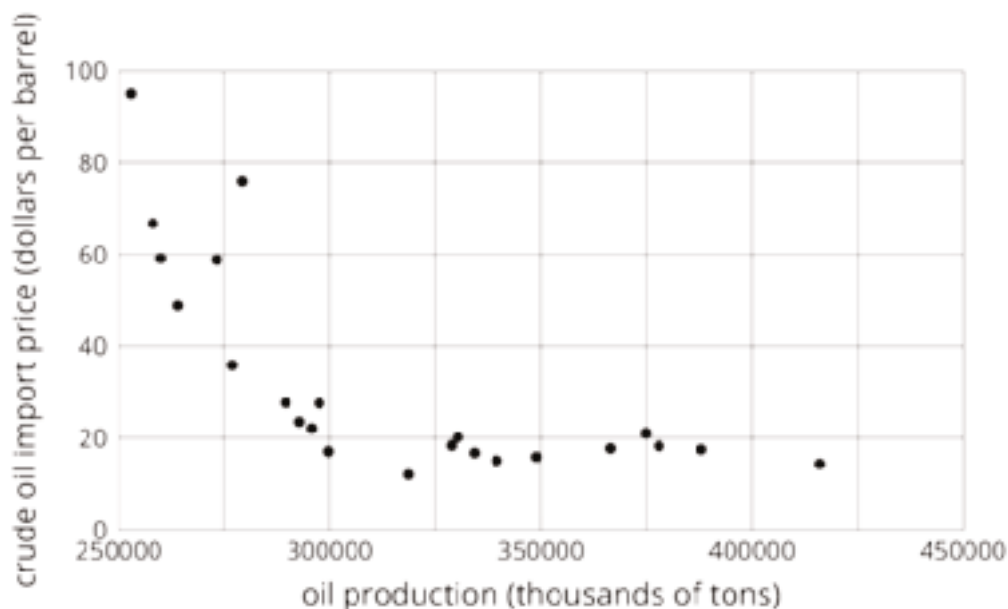
Launch

Display the scatter plot for all to see. Tell students that oil production and price are shown. Production is measured in “thousand tons,” so a point on the left side of this graph represents 250,000 thousand tons (or 250,000,000 tons). Price is measured in “dollars per barrel.” A barrel of oil is equivalent to 42 gallons. The prices in this chart are for crude oil which would still need to be processed to be made useful for gasoline, plastics, or other oil-based materials.

Tell students to signal when they have at least one thing they notice or one thing they wonder. Give students 1 minute of quiet think time followed by whole-class discussion.

Student Task Statement

What do you notice? What do you wonder?



Student Response

Things students may notice:

- The points don't all lie near a line like they have in previous scatter plots.
- As production increases, the price decreases.

Things students may wonder:

- What does "oil production" or "import prices" mean (if the axis labels are unfamiliar)?
- Is it okay to model this relationship using a line?
- Why does it seem to level out at the end?

Activity Synthesis

Ask students to share things they noticed and wondered. Record and display their responses for all to see.

To help focus students on the non-linearity of the association, ask "How is this scatter plot different from ones we have seen so far?" (The data appear to be grouped along a curve and not a straight line.)

Tell students that, even when a linear association seems to be present, it may only fit the data very close to the data that is present. For example, if we cover up the scatter plot to the right of 300,000, there may appear to be a linear association that would fit the data, but it would not apply to the data on the right side of the scatter plot. Additionally, in this particular situation, a linear model

does not make sense in the long term, since a linear model would eventually have a value of zero and then become negative, something that would probably not happen to the price of oil.

7.2 Scatter Plot City

20 minutes

In previous lessons, data was fit with a linear model, but is not appropriate in all cases. In this activity, students individually sort cards of different scatter plots then discuss with a partner. They then re-sort the cards based on linearity and then again by positive or negative association (MP4).

Addressing

- 8.SP.A.1

Instructional Routines

- MLR8: Discussion Supports
- Take Turns

Launch

Arrange students in groups of 2, and give each student a set of cards. Give 5 minutes of quiet work time for the first question followed by partner discussion. For the final two questions, ask students to take turns sorting the cards by putting one card into one of the categories, then explaining their reasoning for putting it there. After both partners have agreed on the placement of that card, the other partner may repeat these steps to sort one more card. Repeat this process until all cards are sorted.

Tell the students that there is a third category for the last two parts of the activity: neither. It is possible for data to have absolutely no association and appear completely random which would not be a linear association or a non-linear association. It is also possible for data to not have a strictly positive or negative association if, for example, the scatter plot seems to increase and then decrease.

Access for Students with Disabilities

Representation: Internalize Comprehension. Chunk this task into more manageable parts to differentiate the degree of difficulty or complexity by beginning with fewer cards. For example, give students a subset of the cards to start with and introduce the remaining cards once students have completed their initial set of matches.

Supports accessibility for: Conceptual processing; Organization

Access for English Language Learners

Speaking: MLR8 Discussion Supports. To support students' explanations for sorting the scatter plots in the way they chose, display sentence frames for students to use when they are working with their partner. For example, "I think ____ because ____ ." or "I (agree/disagree) because ____ ."

Design Principle(s): Support sense-making; Optimize output for (explanation)

Student Task Statement

Your teacher will give you a set of cards. Each card shows a scatter plot.

1. Sort the cards into categories and describe each category.
2. Explain the reasoning behind your categories to your partner. Listen to your partner's reasoning for their categories.
3. Sort the cards into two categories: positive associations and negative associations. Compare your sorting with your partner's and discuss any disagreements.
4. Sort the cards into two categories: linear associations and non-linear associations. Compare your sorting with your partner's and discuss any disagreements.

Student Response

1. Answers vary. Sample categories: decimals and whole numbers, positive and negative associations, more and less variable
2. Answers vary.
3. Positive associations: F and G. Negative associations: B, C, D, and E. Neither: A and H
4. Linear associations: C and E. Non-linear associations: A, B, D, F, G, and H

Activity Synthesis

To help clarify the meaning of linear and non-linear associations as well as positive and negative associations, consider asking some of the following questions:

- "How did you originally group the cards?"
- "Which graphs did you sort into the 'positive association' group?" (F and G)
- "Are there any cards that did not fit with either positive or negative association?" (Yes, the points for card A form an arc that first goes up and then down.)
- "Which graphs did you sort into the 'linear association' group?" (C and E)
- "If a scatter plot has no association at all (neither linear nor non-linear), what might its scatter plot look like?" (A lot of random dots scattered all around the plot with no obvious trend.)

7.3 Clustering

10 minutes

In addition to the type of model appropriate for the data represented in a scatter plot and the positive or negative association, there are other patterns that can be worth noticing before performing a more in-depth analysis. In this activity, students compare scatter plots that have more than one pattern embedded in them. In the discussion following the activity, this feature is identified as *clustering*.

As students compare the scatter plots in this activity, monitor for students who notice clustering of the data and name it in their own words.

Addressing

- 8.SP.A.1

Instructional Routines

- MLR2: Collect and Display
- Think Pair Share

Launch

Keep students in groups of 2. Allow 2 minutes quiet think time followed by partner and whole-class discussion.

Access for Students with Disabilities

Engagement: Develop Effort and Persistence. Encourage and support opportunities for peer interactions. Display sentence frames to support student conversation such as: “____ and ____ are different because...”, “One thing that is the same is...”, and “What do ____ and ____ have in common?”

Supports accessibility for: Language; Social-emotional skills

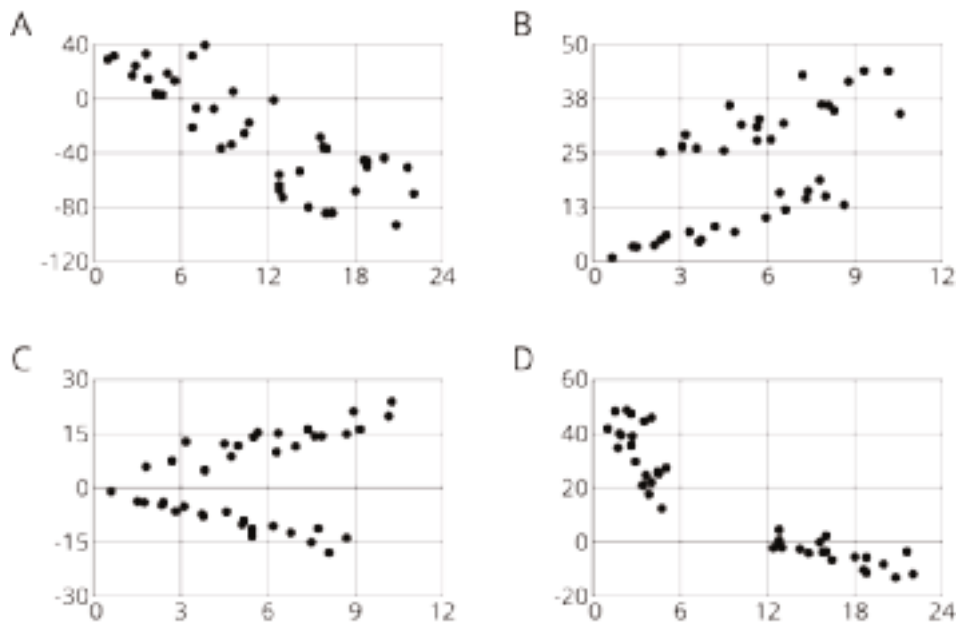
Access for English Language Learners

Conversing, Representing, Writing: MLR2 Collect and Display. As students discuss how the scatter plots are alike, and how they are different, capture the vocabulary and phrases students use to describe the patterns they notice. Listen for students who refer to clustering, or multiple patterns in one graph. Record their language and any relevant sketches on a visual display that can be referenced in future discussions. This will help students produce and make sense of the language needed to communicate about the relationship between quantities represented graphically.

Design Principle(s): Support sense-making; Maximize meta-awareness

Student Task Statement

How are these scatter plots alike? How are they different?



Student Response

Answers vary. Sample responses:

- Plots A and D both have negative associations, but D has a gap in the data.
- Plots B and C both have 2 groups of data that would have different lines of fit. In B, both lines of fit would have about the same slope, but in C, one slope is positive and the other is negative.
- Plots A, C, and D each have data in quadrants I and IV, but Plot B only has data in Quadrant I.

Activity Synthesis

The purpose of this discussion is for students to understand what it means for data to appear in clusters and recognize when they might appear.

Select previously identified students to share their observations about the scatter plots. Tell students that when data seems to have more than one pattern, it is called *clustering*. Clustering of the data like in graph B, C, and D can reveal hidden patterns. Usually, clustering means there are subgroups within our data that may represent different trends.

For example, in Plot B, the data may represent body measurements of a certain species of bird. Although the data originally came from a group that made sense (a single species), there appear to be subgroups that have a large influence on the data as well. The lower half of the data may represent females and the upper half may represent males, so we can see that there are different patterns within the different subgroups.

When clustering is present, it may be helpful to investigate the cause of the separation and analyze the data within the subgroups rather than as a whole.

Lesson Synthesis

To help summarize the lesson, ask:

- “In your own words, what is a non-linear association?” (The points in a scatter plot do not lie along a straight line.)
- “What does a non-linear association mean?” (There is not a constant increase or decrease of one variable based on the other.)
- “In your own words, what are clusters in data?” (The points in a scatter plot are clumped together in different groups.)
- “What do clusters usually mean?” (There may be multiple patterns present within the data. Perhaps there are subgroups that show different patterns.)

7.4 Make Your Own Scatter Plot

Cool Down: 5 minutes

Students show their understanding of positive and negative association, linear and non-linear association, and clustering of data by drawing scatter plots with these conditions (MP7).

Addressing

- 8.SP.A.1

Launch

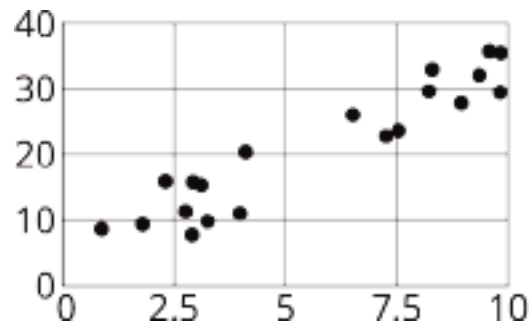
Advise students to use around 20 data points in their graph. This is enough to show the intention of the graph without being tedious.

Student Task Statement

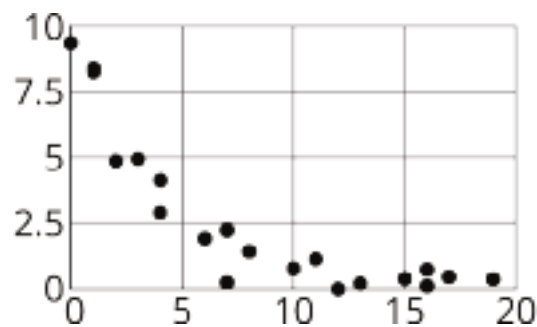
1. Draw a scatter plot that shows a positive linear association and clustering.
2. Draw a scatter plot that shows a negative non-linear association and no clustering.

Student Response

1. Answers vary. Sample response:

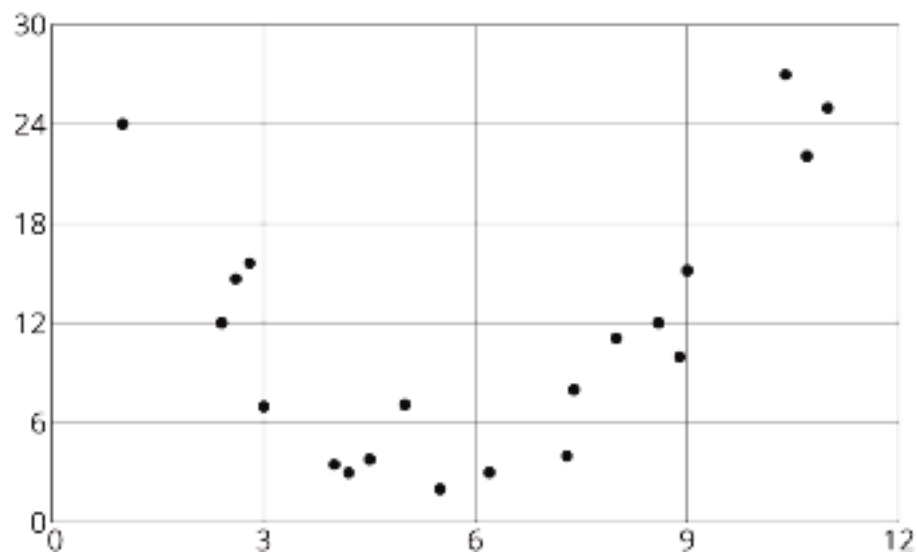


2. Answers vary. Sample response:



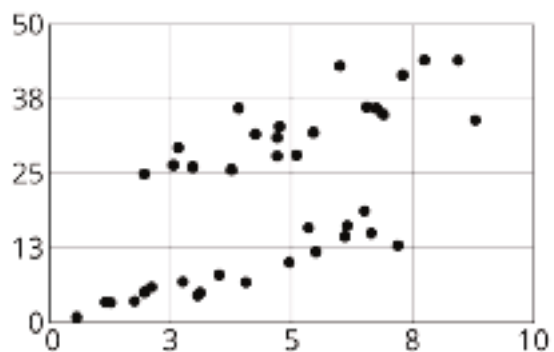
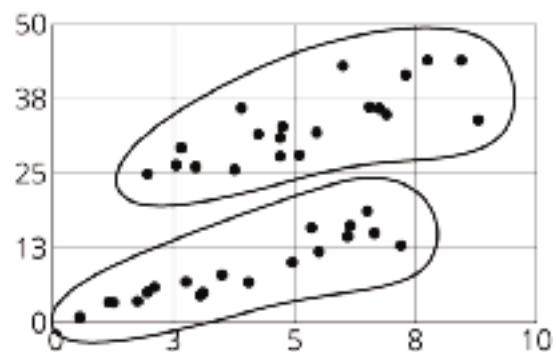
Student Lesson Summary

Sometimes a scatter plot shows an association that is *not* linear:



We call such an association a *non-linear association*. In later grades, you will study functions that can be models for non-linear associations.

Sometimes in a scatter plot we can see separate groups of points.



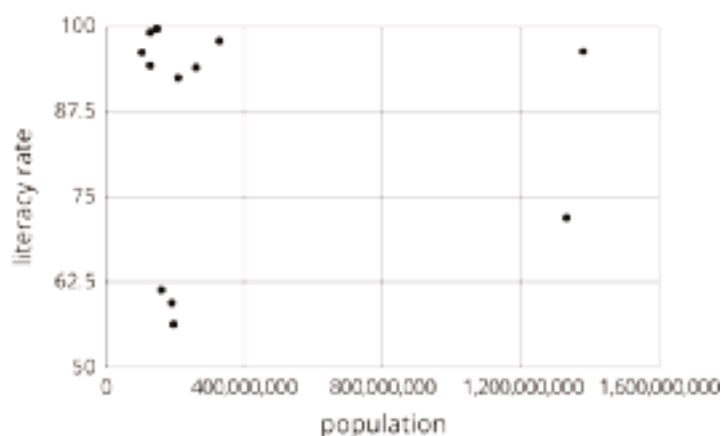
We call these groups *clusters*.

Lesson 7 Practice Problems

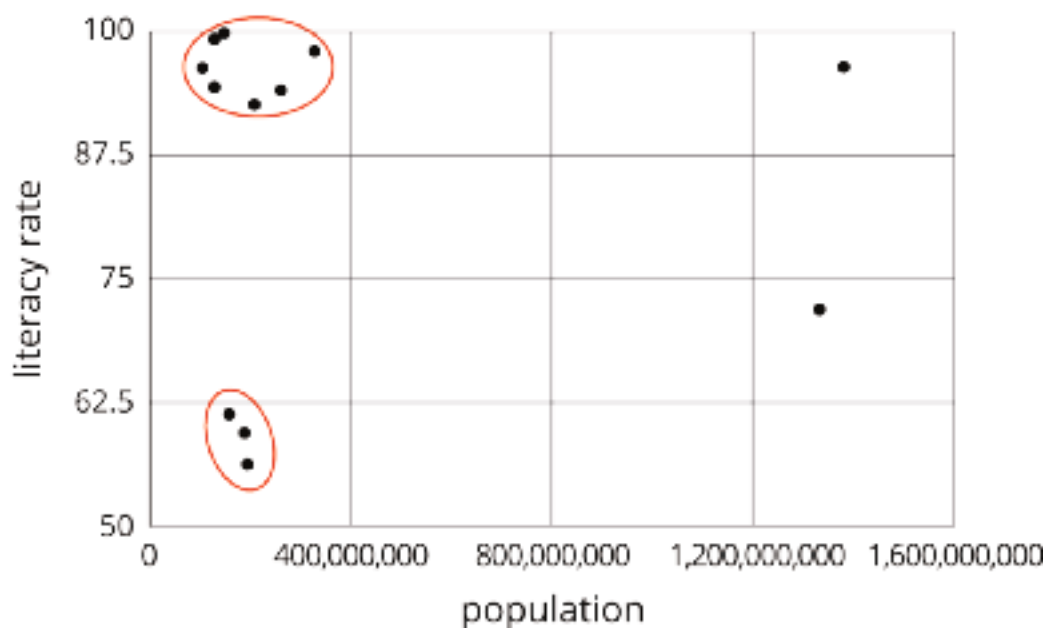
Problem 1

Statement

Literacy rate and population for the 12 countries with more than 100 million people are shown in the scatter plot. Circle any clusters in the data.



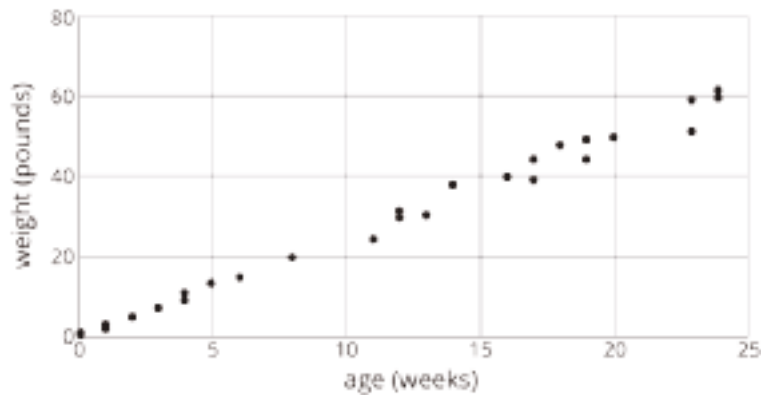
Solution



Problem 2

Statement

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

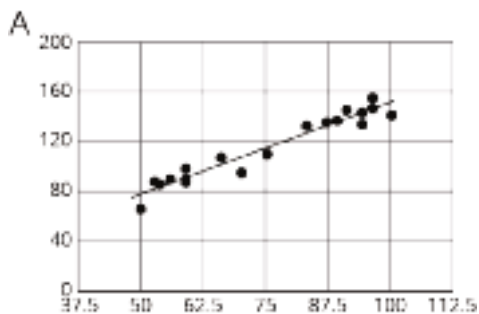
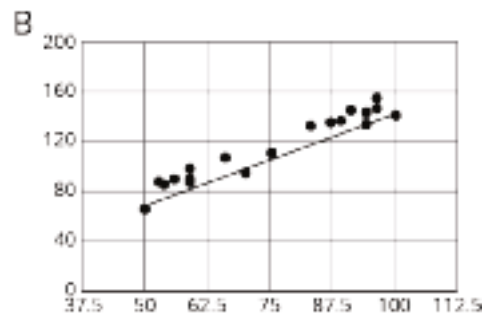
Solution

["A", "C"]

Problem 3

Statement

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



Solution

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.

(From Unit 6, Lesson 5.)

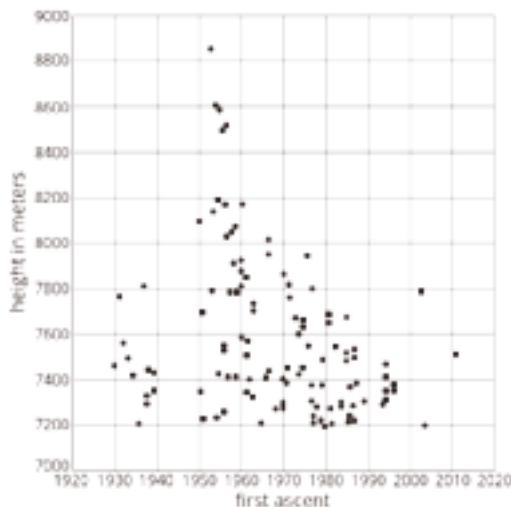
Problem 4

Statement

Here is a scatter plot of data for some of the tallest mountains on Earth.

The heights in meters and year of first recorded ascent is shown. Mount Everest is the tallest mountain in this set of data.

- Estimate the height of Mount Everest.
- Estimate the year of the first recorded ascent of Mount Everest.



Solution

- Approximately 8,848 meters (The vertical coordinate of the data point with the greatest vertical value is closer to 8800 than 9000.)
- Approximately 1953 (The horizontal coordinate of the same point is slightly to the right of 1950.)

(From Unit 6, Lesson 3.)

Problem 5

Statement

A cone has a volume V , radius r , and a height of 12 cm.

- A cone has the same height and $\frac{1}{3}$ of the radius of the original cone. Write an expression for its volume.
- A cone has the same height and 3 times the radius of the original cone. Write an expression for its volume.

Solution

- $\frac{V}{9}$
- $9V$

(From Unit 5, Lesson 18.)

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 Targets

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

Lesson Narrative

In this lesson, students bring everything they have studied in the unit so far to analyze and interpret bivariate data in context (MP4). They create a scatter plot, identify outliers, fit a line, and determine and interpret the slope of the line. They compare actual and predicted values. They reflect on what they have learned about modeling bivariate data.

Alignments

Addressing

- 8.SP.A.1: 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.
- 8.SP.A.2: 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.
- 8.SP.A.3: Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope 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

- 8.SP.A.3: Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope 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

- MLR1: Stronger and Clearer Each Time
- MLR8: Discussion Supports

- Think Pair Share

Required Materials

Dried linguine pasta

We specified linguine since it is flatter and less likely to roll around than spaghetti.

straightedge, but sometimes it is preferable to use an unruled straightedge, like a blank index card.

Straightedges

A rigid edge that can be used for drawing line segments. Sometimes a ruler is okay to use as a

Required Preparation

Class data (or data from another group if it fits the classroom culture better) from the second lesson in the unit on arm span and height.

Student Learning Goals

Let's analyze data like a pro.

8.1 Speed vs. Step Length

Warm Up: 5 minutes

The purpose of this warm-up is for students to interpret the rate of change of a function to describe a trend (MP2). 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.

Building Towards

- 8.SP.A.3

Launch

Give students 2 minutes of quiet work time followed by a whole-class discussion.

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?

Student Response

For every one 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?”

8.2 Animal Brains

15 minutes (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 (MP2).

Addressing

- 8.SP.A.1
- 8.SP.A.2
- 8.SP.A.3

Instructional Routines

- MLR8: Discussion Supports
- Think Pair Share

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.

If using the digital activity, students can still work independently to analyze the scatter plot and answer the prompts. Then students can discuss their thinking in groups of 2 followed by whole-class discussion.

Anticipated Misconceptions

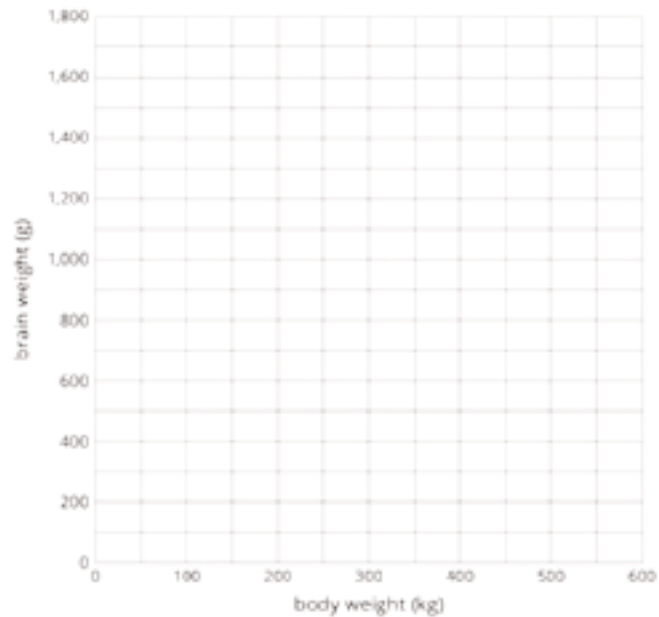
When estimating slope, some students won't use the scales of the axes correctly, so the slope is reported incorrectly. Some students may not notice the different units of weight used on each axis.

Student Task Statement

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

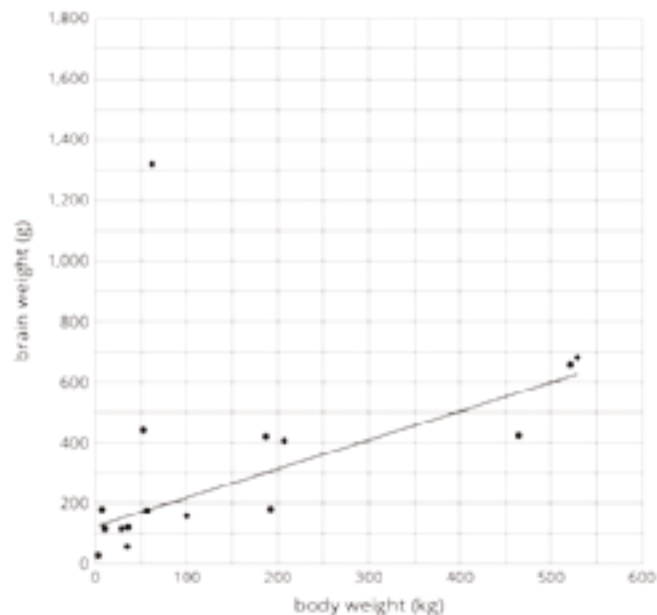
Use the data in the table to make a scatter plot. Are there any outliers?

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
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



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 Response



The human data is an outlier.

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

2. 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. The chimpanzee data may also be an outlier.

Are You Ready for More?

Use one of the suggestions or find another set of data that interested 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 grossed 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 two variables? Explain your reasoning.

Student Response

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?"
- "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)

Access for English Language Learners

Speaking: MLR8 Discussion Supports. Use this routine to support whole-class discussion. For each response or observation that is shared, ask students to restate and/or revoice what they heard using mathematical language. Consider providing students time to restate what they hear to a partner, before selecting one or two students to share with the class. Ask the original speaker if their peer was accurately able to restate their thinking. Call students' attention to any words or phrases that helped to clarify the original statement.

Design Principle(s): Support sense-making

8.3 Equal Body Dimensions

15 minutes (there is a digital version of this activity)

In this activity students create another scatter plot to analyze the data they collected about their classmates in a previous lesson (MP4). 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 $y = x$ which is symmetric when switching which variable is represented on each axis. If possible, identify any groups who have axes switched to bring up in the discussion.

Note: Some students may be sensitive about their body measurements and providing alternate data allows the class to work with actual values without making students uncomfortable. Depending on your class, consider providing a similar data set to the one collected in the earlier lesson (measurements from the staff, a different class, or invented data that is similar to the data collected).

Addressing

- 8.SP.A.1
- 8.SP.A.2

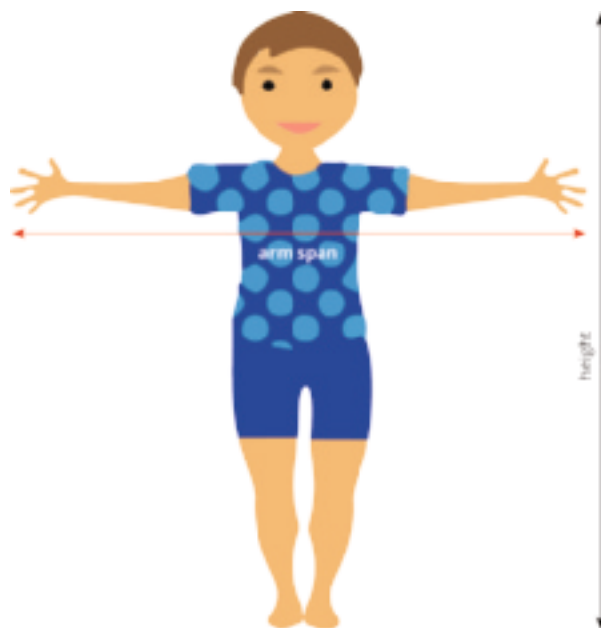
Instructional Routines

- MLR1: Stronger and Clearer Each Time

Launch

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

Display the class data collected earlier in the unit. Display this image to remind students what their data represents.



If using the digital activity, students will create their scatter plots using technology.

Access for Students with Disabilities

Action and Expression: Provide Access for Physical Action. Provide access to the digital version of the task or graphing software. If needed, provide students with a completed table of student data of arm span and height collected in an earlier lesson.

Supports accessibility for: Visual-spatial processing; Conceptual processing; Organization

Access for English Language Learners

Writing, Speaking, Listening: MLR1 Stronger and Clearer Each Time. Use this routine to give students a structured opportunity to revise and refine their 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.” Ask each student to meet with 2–3 other partners in a row for feedback. Provide students with prompts for feedback that will help them strengthen their ideas and clarify their language (e.g., “Why do you think....?”, “Can you give an example?”, “How did you determine a better line?”, etc.). Students can borrow ideas and language from each partner to strengthen their final version.

Design Principle(s): Optimize output (for generalization)

Student Task Statement

Earlier, your class gathered data on height and arm span.

1. Sometimes a person’s arm span is the same as their height. Is this true for anyone in the class?

2. Make a scatter plot for the arm span and height data, and describe any association.



3. Is the line $y = x$ a good fit for the data? If so, explain why. If not, find the equation of a line that fits the data better.
4. Examine the scatter plot. Which person in your class has the *largest* ratio between their arm span and their height? Explain or show your reasoning.

Student Response

Answers vary based on class data.

1. Any data points on the line $y = x$ represent students who have equal arm spans and heights.
2. Graphs should have one axis labeled with height (cm) and one with arm span (cm).
There is a positive, linear association between height and arm span, because as height increases, arm span tends to increase.
3. The line $y = x$ may be a good fit, because in general, human arm span and height are about the same.
4. The person with the largest ratio will have a data point that is farther from the line $y = x$ than the data points of the other students.

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 fit the data?” (Answers vary based on class data, but should be close to 1.)
- If any groups had axes switched, select these groups to show their scatter plots. “What would a point on the line $y = x$ represent in each graph?” (A person whose arm span and height are identical. It does not matter which way the axes are drawn for this line.)
- “What does it mean about a person whose point is above the line $y = x$? What does it mean about a person whose point is below the line?” (Depending on the axes, a person whose point is off of the line has a longer arm span than expected for their height or is taller than expected for their arm span.)
- “Suppose you measure height and arm span for people in you neighborhood. When you make a scatter plot, you notice two clusters of data: one group of data in the lower left of the scatter plot and another group in the upper right. What does this mean? Why do you think there may be these clusters?” (It may mean there are a lot of small children and taller adults.)

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?” (Two 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.)

8.4 Drawing a Line

Cool Down: 5 minutes

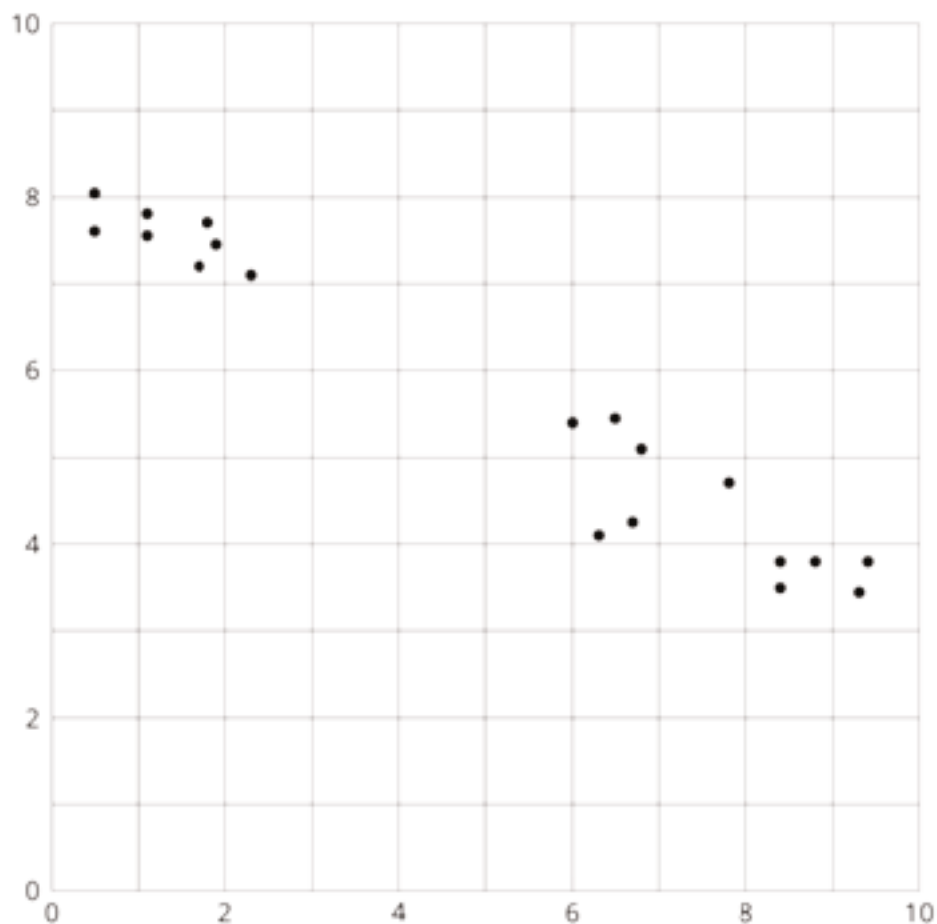
Students use a given scatter plot to propose a good linear model for the data. They then use that model to predict the value of a new data point within the data.

Addressing

- 8.SP.A.2
- 8.SP.A.3

Student Task Statement

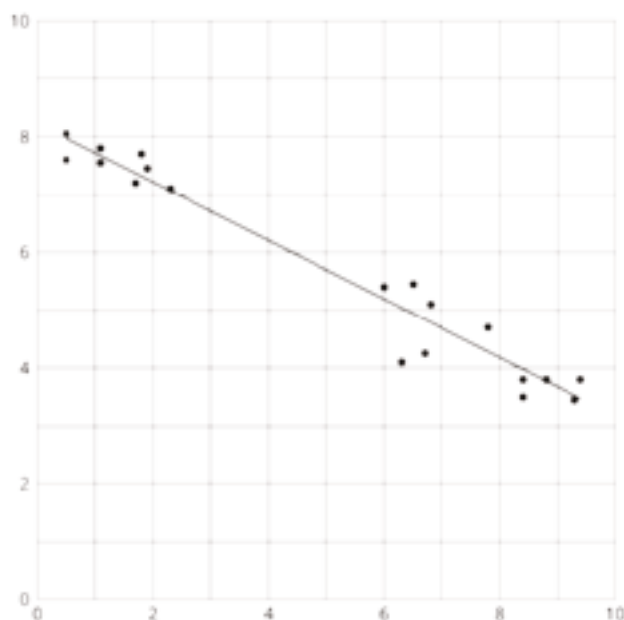
1. Draw a line on the scatter plot that fits the data well.



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?

Student Response

1. Answers vary. Sample response:



2. Answers vary. Sample responses: 6, 6.1, 5.9

Student 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, and 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.

Lesson 8 Practice Problems

Problem 1

Statement

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.

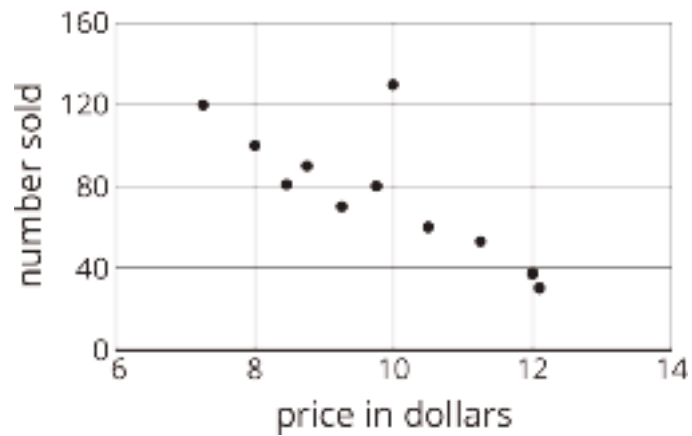
price in dollars	number sold
11.25	53
10.50	60
12.10	30
8.45	81
9.25	70
9.75	80
7.25	120
12	37
9.99	130
7.99	100
8.75	90



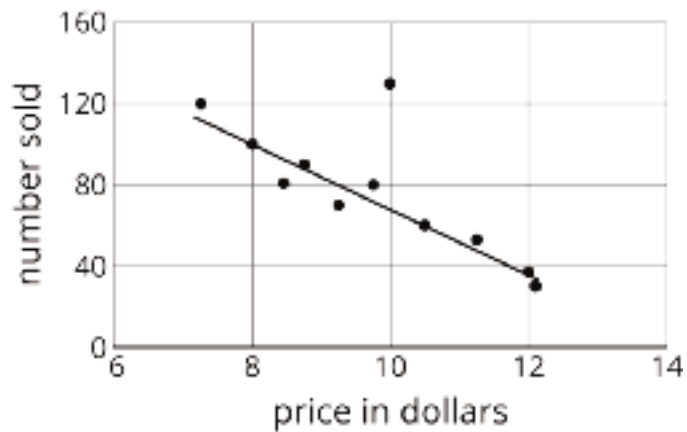
- Draw a scatter plot of this data. Label the axes.
- Are there any outliers? Explain your reasoning.
- If there is a relationship between the variables, explain what it is.
- Remove any outliers, and draw a line that you think is a good fit for the data.

Solution

a.



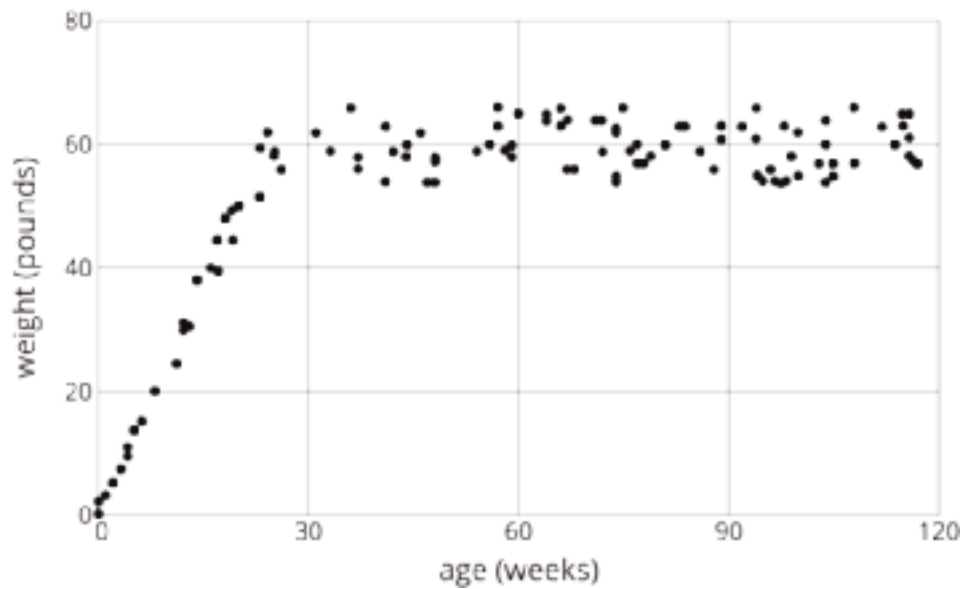
- b. Yes, at (9.99, 130). This point is much higher than expected on the scatter plot.
- c. There is a negative linear relationship between the variables. When the price increases, the number of books sold decreases.
- d. Answers vary. Sample response:



Problem 2

Statement

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

Solution

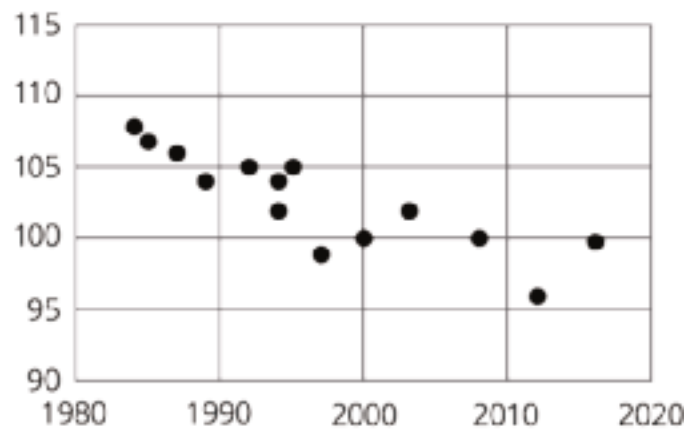
["B", "C"]

(From Unit 6, Lesson 7.)

Problem 3

Statement

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



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

Solution

C

(From Unit 6, Lesson 6.)

Section: Associations in Categorical Data

Lesson 9: Looking for Associations

Goals

- Calculate relative frequencies, and describe (orally and in writing) associations between variables using a relative frequency table.
- Coordinate (orally and in writing) two-way tables, bar graphs, and segmented bar graphs representing the same data.

Learning Targets

- I can identify the same data represented in a bar graph, a segmented bar graph, and a two-way table.
- I can use a two-way frequency table or relative frequency table to find associations among variables.

Lesson Narrative

In this lesson, students study categorical data displayed in two-way tables, bar graphs, and segmented bar graphs. The different graphical representations help students visualize the frequencies and relative frequencies, aiding them in making judgement about whether there is evidence of an association or not in the next lesson. While the concepts and structures in this lesson are not very complex, there are many new terms and representations, and students are given the opportunity to study them carefully so that they can make sense of them (MP1).

Alignments

Addressing

- 8.SP.A.4: Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?

Building Towards

- 8.SP.A.4: Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. For example, collect data from students in your class

on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?

Instructional Routines

- MLR2: Collect and Display
- MLR8: Discussion Supports
- Notice and Wonder
- Think Pair Share

Required Materials

Pre-printed cards, cut from copies of the Instructional master

Required Preparation

Print and cut up cards from the Matching Representations Instructional master. Prepare 1 set of cards for every 2 students.

Student Learning Goals

Let's look for associations in data.

9.1 Notice and Wonder: Bar Association

Warm Up: 5 minutes

The purpose of this warm-up is for students to become familiar with a bar graph by noticing and wondering things about it. While reading a bar graph is a review of a previous grade's work, it is an important for students to look for patterns of association in categorical data.

Building Towards

- 8.SP.A.4

Instructional Routines

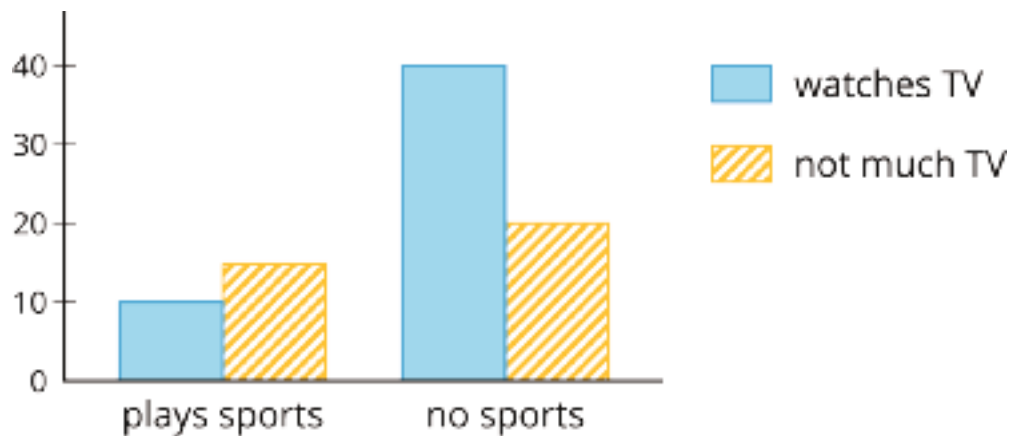
- Notice and Wonder

Launch

Tell students you are going to show them a graph and that their job is to find at least one thing they notice and one thing they wonder about the graph. Give students 1 minute of quiet think time followed by a whole-class discussion.

Student Task Statement

What do you notice? What do you wonder?



Student Response

Answers vary.

Students may notice:

- More people do not play sports than do.
- More people watch TV than watch not much TV.
- The red bar is taller than the blue for “plays sports” but the other way around for “no sports.”

Students may wonder:

- What does “watches TV” mean?
- What does “not much TV” mean?
- Who is represented in the poll?
- Can a person who watches some TV be in the blue and yellow bars?

Activity Synthesis

Ask students to share things they noticed and wondered. Record and display their responses for all to see. If students wondered who the data represents, tell them the data was collected from a survey of students. They were asked whether they played a sport or not and whether they watched more or less than 1 hour of TV each night.

If they did not wonder these things, tell them this information and then ask, “Based on the graph, do you think playing a sport and the amount of time watching TV are related?” (Yes. Among people who watch more than an hour of TV, there are a lot more who do not play a sport than people who do. For people who play a sport, more do not watch TV than do and this is reversed for those who do not play a sport.)

9.2 Card Sort: Matching Representations

20 minutes

In this activity students become familiar with two-way tables, clustered bar graphs, and segmented bar graphs by matching different situations. They label the diagrams to match the data given and create a table to match the data shown in one of the bar graphs.

Addressing

- 8.SP.A.4

Instructional Routines

- MLR2: Collect and Display
- MLR8: Discussion Supports

Launch

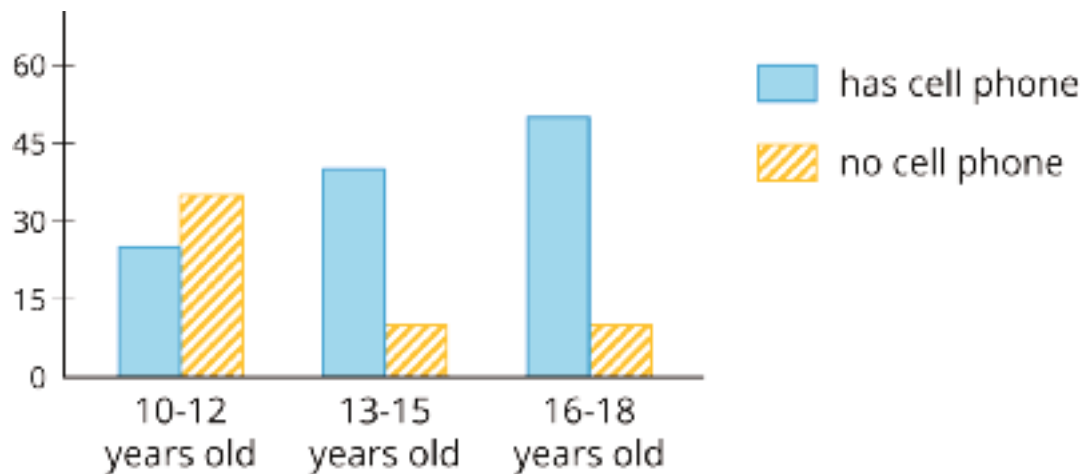
Arrange students in groups of 2–3. Display the two-way table from the activity for all to see. Use MLR 2 (Collect and Display) to gather student language used to describe two-way tables, bar graphs, and segmented bar graphs to be referenced and revised throughout later activities.

	has cell phone	does not have cell phone	total
10 to 12 years old	25	35	60
13 to 15 years old	40	10	50
16 to 18 years old	50	10	60
total	115	55	170

Ask students:

- “Why do you think this is called a ‘two-way table?’” (The table represents 2 variables. You can think of the people who were surveyed either by age group or by whether or not the person has a cell phone.)
- “What does the 25 mean?” (25 people 10 to 12 years old have cell phones.)
- “What does the 55 mean?” (55 of the people who were surveyed do not have a cell phone.)
- “How many people are represented in this table?” (170 people.)

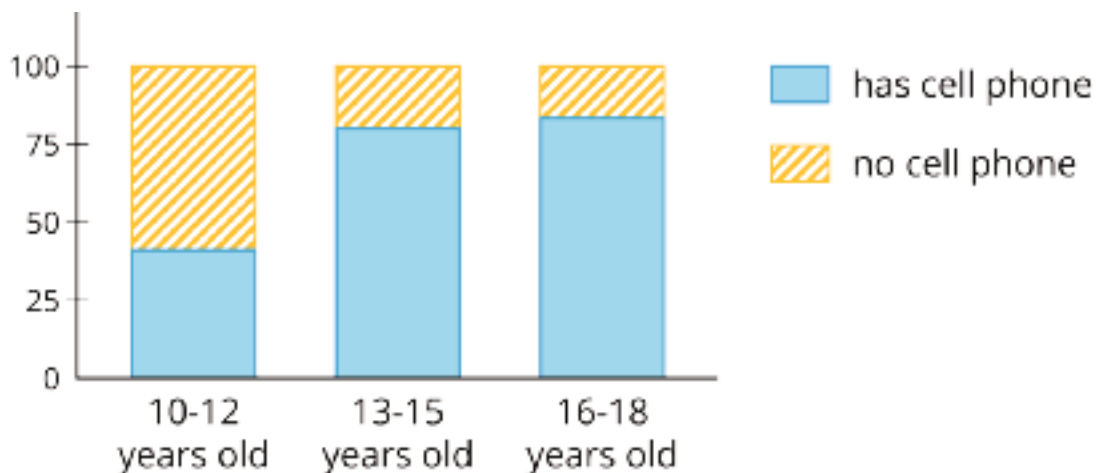
Display the bar graph from the activity for all to see.



Ask students:

- “Why do you think this is called a ‘bar graph?’” (Each bar represents how many people are in each group named on the bottom.)
- “What do the blue bars represent? How do you know?” (The blue bars represent how many people in each group have cell phones. I can tell by the key on the right.)
- “Where does the 25 from the table show up in this bar graph?” (The first blue bar shows that 25 10–12 year olds have cell phones.)

Display the **segmented bar graph** from the activity for all to see.



Ask students:

- “Why do you think this is called a ‘segmented bar graph?’” (Each bar is segmented into pieces.)
- “What do you think the vertical axis represents?” (The percentage in each section.)

Distribute 1 set of the pre-cut cards from the Instructional master to each group.

Access for Students with Disabilities

Representation: Internalize Comprehension. Chunk this task into more manageable parts to differentiate the degree of difficulty or complexity by beginning with fewer cards. For example, give students a subset of the cards to start with and introduce the remaining cards once students have completed their initial set of matches.

Supports accessibility for: Conceptual processing; Organization

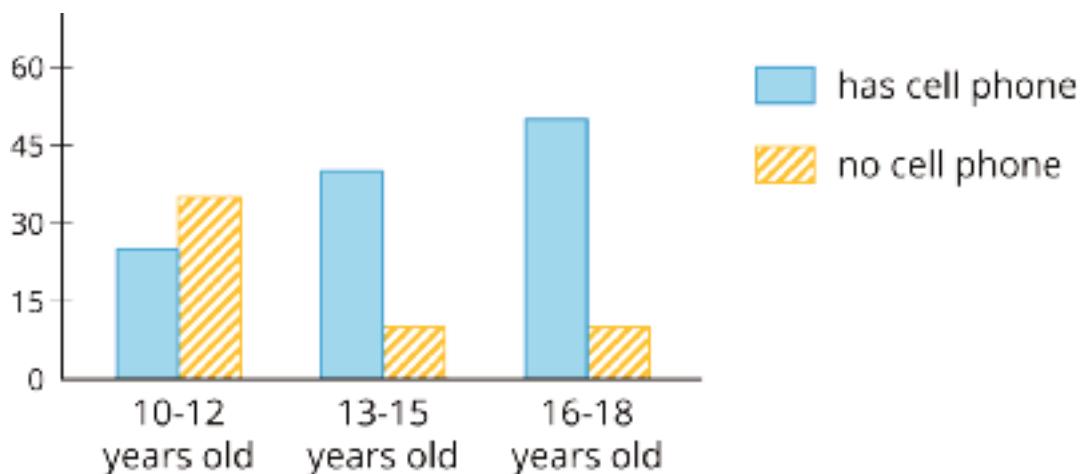
Student Task Statement

Your teacher will hand out some cards.

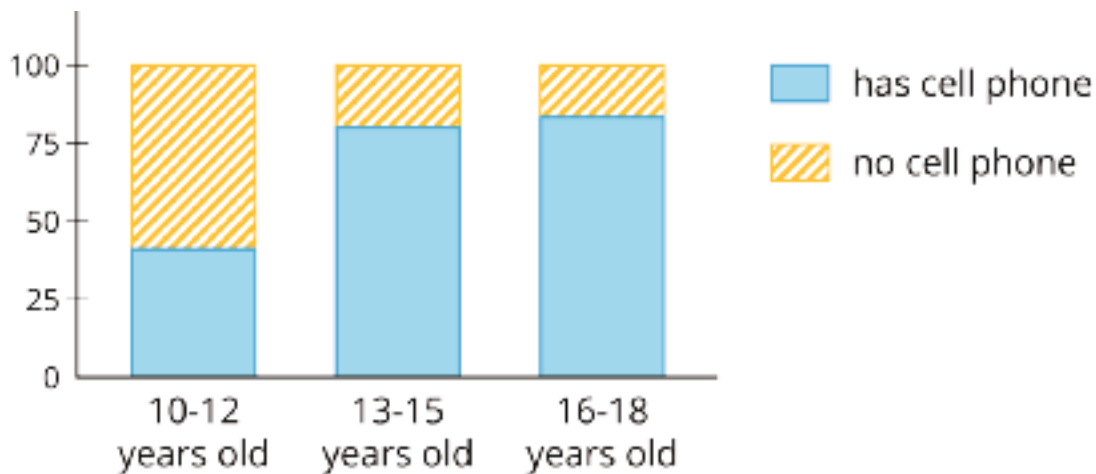
Some cards show **two-way tables** like this:

	has cell phone	does not have cell phone	total
10 to 12 years old	25	35	60
13 to 15 years old	40	10	50
16 to 18 years old	50	10	60
total	115	55	170

Some cards show bar graphs like this:



Some cards show **segmented bar graphs** like this:



The bar graphs and segmented bar graphs have their labels removed.

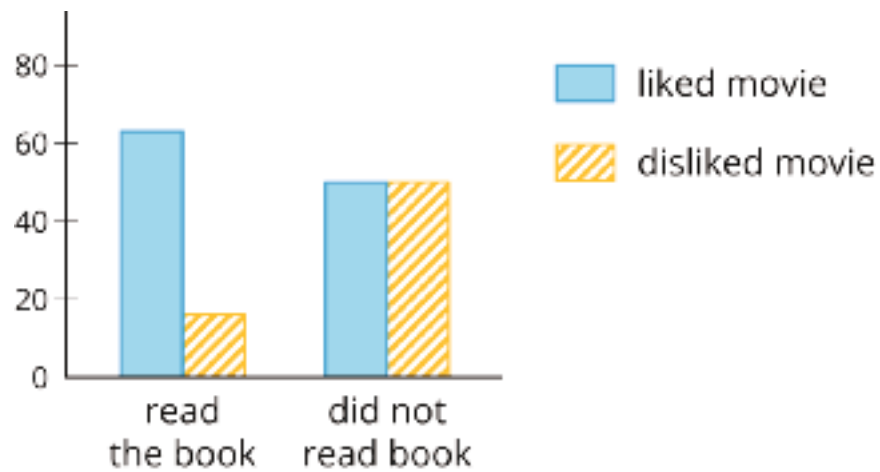
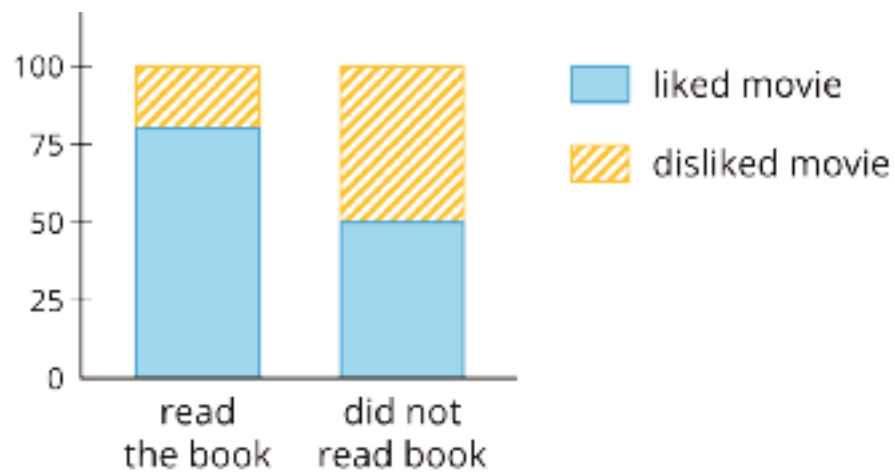
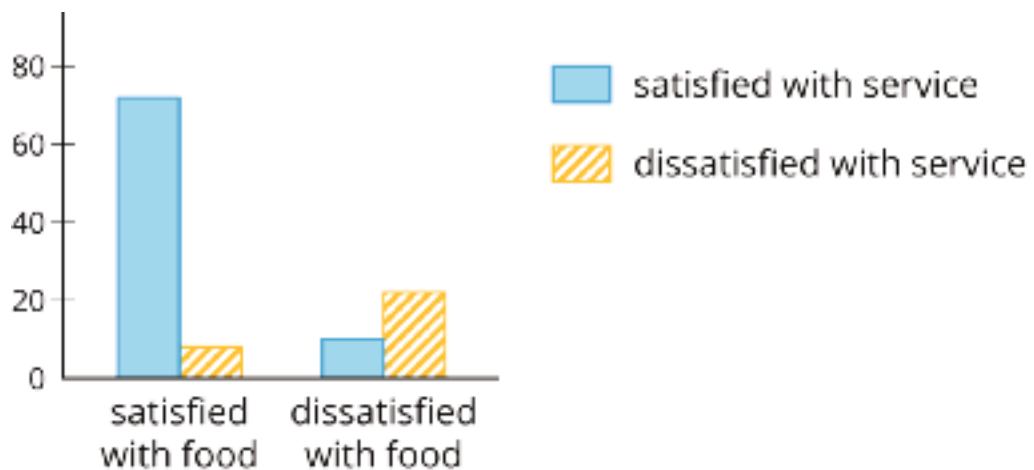
1. Put all the cards that describe the same situation in the same group.
2. One of the groups does not have a two-way table. Make a two-way table for the situation described by the graphs in the group.
3. Label the bar graphs and segmented bar graphs so that the categories represented by each bar are indicated.
4. Describe in your own words the kind of information shown by a segmented bar graph.

Student Response

1. See Instructional master for the groupings.
2. Answers vary. Sample response:

	ate breakfast	skipped breakfast	total
male	48	12	60
female	48	16	64
total	96	28	124

3. Answers vary. Possible response:

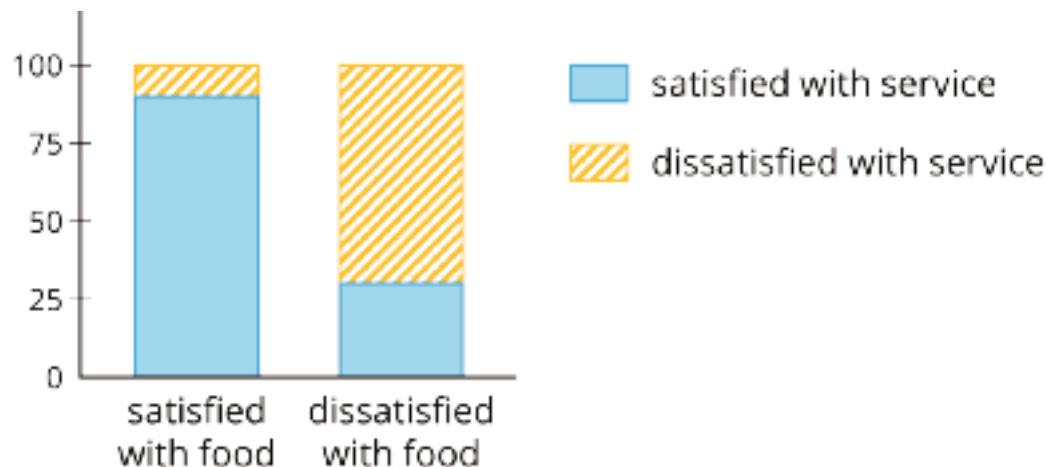


4. In a segmented bar graph, each bar represents 100% of a category. The bar is split into different subcategories that are stacked on top of each other so you can compare their percentages.

Are You Ready for More?

One of the segmented bar graphs is missing. Construct a segmented bar graph that matches the other representations.

Student Response



Activity Synthesis

The purpose of this activity is to help students understand the connections among the three representations of data. It also helps students see the importance of labeled visual representations.

To highlight the connections, ask:

- “What were some strategies you used to match the cards?” (The book table had the same number for those who liked the movie and those who did not like the movie among those who did not read the book, so I knew the two graphs would show equal sized bars.)
- “Were some cards easier to match than others? Explain.” (The cards showing the actual numbers were easier to match than the cards showing the percentages.)
- “What strategy did you use to create the two-way data table?” (I started by thinking of two yes or no questions that could be asked and use those to label the rows and columns of the table.)

Access for English Language Learners

Speaking: MLR8 Discussion Supports. Use this routine to support whole-class discussion. For each strategy that is shared, ask students to restate and/or revoice what they heard using mathematical language. Consider providing students time to restate what they hear to a partner, before selecting one or two students to share with the class. Ask the original speaker if their peer was accurately able to restate their thinking. Call students' attention to any words or phrases that helped to clarify the original statement. This will provide more students with an opportunity to produce language as they interpret the reasoning of others.

Design Principle(s): Support sense-making

9.3 Building Another Type of Two-Way Table

10 minutes

In this activity, students create two-way tables displaying relative frequency. The relative frequency table converts the actual frequency data to percentages which can be useful when comparing groups that include different totals. Finally, students use the relative frequencies to look for a pattern in the data. In the following lesson, students will work with associations in categorical data more explicitly. In this activity, students should use an informal understanding of association to think about whether one is present in the data based on the relative frequency table.

Addressing

- 8.SP.A.4

Instructional Routines

- MLR2: Collect and Display
- Think Pair Share

Launch

Keep students in groups of 2–3. After an introduction to relative frequency tables, allow students 3 minutes quiet work time followed by partner discussion and whole-class discussion.

Display the table for all to see.

	watch the news daily	does not watch the news daily	total
younger than 18	30	80	110
18 or older	10	5	15
total	40	85	125

Ask students, "Based on this data, who is more likely to watch the news daily: someone who is younger than 18 or someone who is 18 or older?"

Tell students that, based on the numbers in the table, there are more younger people who watch the news (30) than older (10). On the other hand, the survey reached out to 110 young people and only 15 older people. Without looking at the whole table, that information may have been missed.

In cases like this, finding a relative frequency including percentages can be more helpful than looking at the actual frequency, which is what they are going to do now.

Use MLR 2 (Collect and Display) to co-create a graphic organizer with students comparing and contrasting the definitions and characteristics of "actual frequency" and "relative frequency" to be referenced and revised throughout this and later activities.

Access for English Language Learners

Conversing, Representing, Writing: MLR2 Collect and Display. As students work in groups, capture the vocabulary and phrases students use when comparing and contrasting the definitions and characteristics of “actual frequency” and “relative frequency.” Listen for students who identify pros and cons for each representation and when one is preferable compared to the other. Record their language on a visual display that can be referenced in future discussions. This routine provides feedback to students that supports sense-making and increases meta-awareness of mathematical language.

Design Principle(s): Support sense-making; Maximize meta-awareness

Anticipated Misconceptions

Students may not calculate relative frequencies correctly. Look to see if they are dividing into the total for each row, instead of some other number in the row, or the total for the entire table.

Student Task Statement

Here is a two-way table that shows data about cell phone usage among children aged 10 to 18.

	has cell phone	does not have cell phone	total
10 to 12 years old	25	35	60
13 to 15 years old	40	10	50
16 to 18 years old	50	10	60
total	115	55	170

1. Complete the table. In each row, the entries for “has cell phone” and “does not have cell phone” should have the total 100%. Round entries to the nearest percentage point.

	has cell phone	does not have cell phone	total
10 to 12 years old	42%		
13 to 15 years old			100%
16 to 18 years old		17%	

This is still a two-way table. Instead of showing *frequency*, this table shows **relative frequency**.

2. Two-way tables that show relative frequencies often don't include a "total" row at the bottom. Why?
3. Is there an association between age and cell phone use? How does the two-way table of relative frequencies help to illustrate this?

Student Response

1.

	has cell phone	does not have cell phone	total
10 to 12 years old	42%	58% since $35 \div 60 \approx 0.58$	100%
13 to 15 years old	80% since $40 \div 50 = 0.8$	20% since $10 \div 50 = 0.2$	100%
16 to 18 years old	83% since $50 \div 60 \approx 0.83$	17%	100%

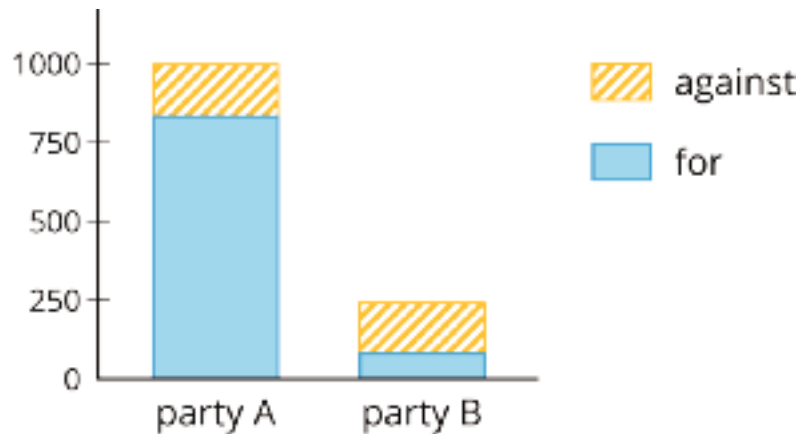
2. The columns are not percentages of the same number, so you can't add them. The percentages in any given column are not likely to add up to 100%.
3. Yes, there is an association between age and cell phone use. A much higher percentage of children from 13 to 18 years old have cell phones than children from 10 to 12 years old do.

Are You Ready for More?

A pollster attends a rally and surveys many of the participants about whether they associate with political Party A or political Party B and whether they are for or against Proposition 3.14 going up for vote soon. The results are sorted into the table shown.

	for	against
party A	832	165
party B	80	160

- A news station reports these results by saying, "A poll shows that about the same number of people from both parties are voting against Proposition 3.14."
- A second news station shows this graphic.

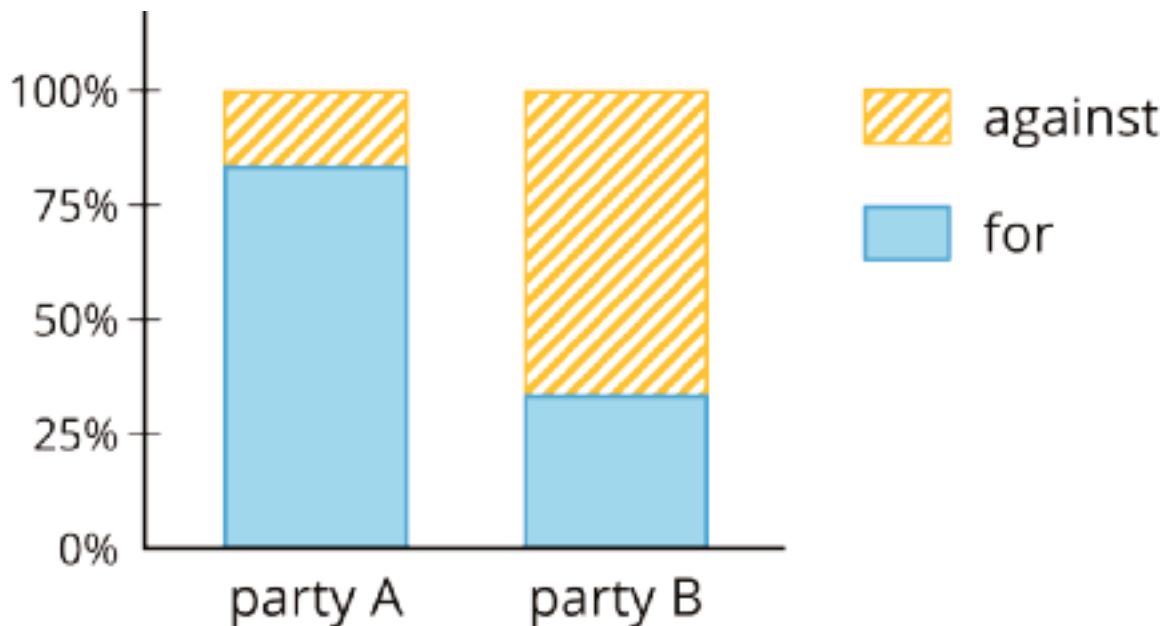


1. Are any of the news reports misleading? Explain your reasoning.
2. Create a headline, graphic, and short description that more accurately represents the data in the table.

Student Response

Answers vary. Sample response:

1. Yes, the first news station did not take into account that there were not as many people from party B that were polled. The graphic from the second news station is hard to understand because the bars from Party B are so much smaller than from Party A, so it is hard to tell what it is trying to say.
2. Headline: Party A is For Proposition 3.14 while Party B is Against



Description: Based on this poll, more than 80% of people in Party A are for Proposition 3.14, while only 33% of people in Party B share the same opinion. There seems to be a big divide along party lines for the proposition!

Activity Synthesis

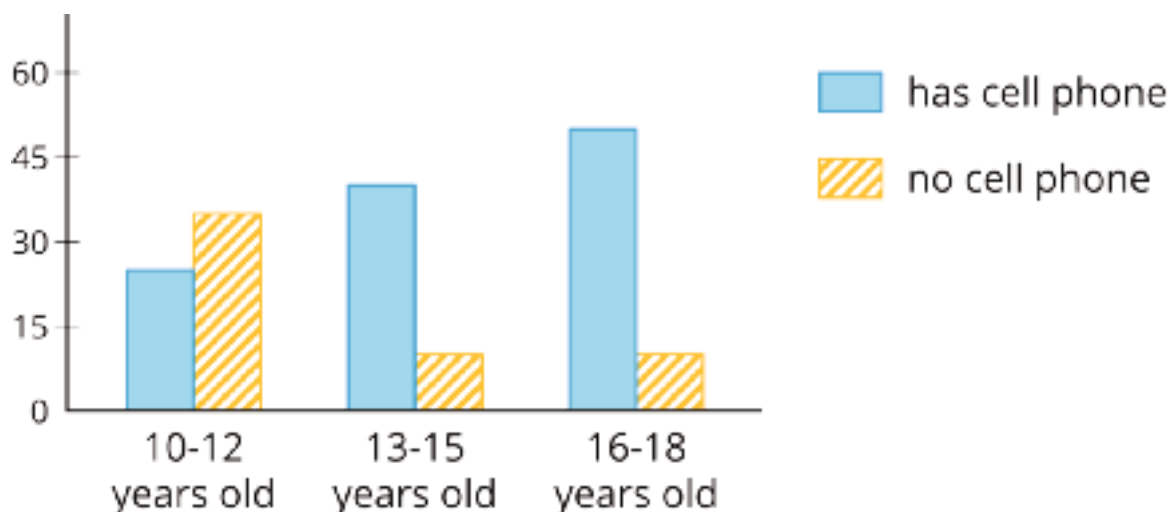
The purpose of this discussion is to help students see the usefulness of two-way tables that display relative frequency.

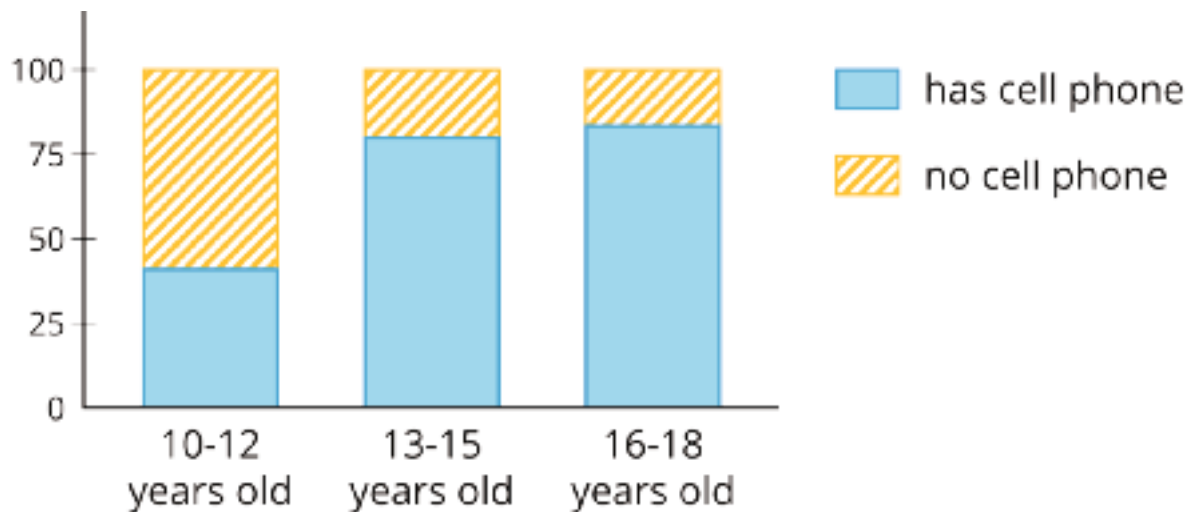
Some questions for discussion:

- “What does the 42% in table mean?” (42% of people 10 to 12 years old have a cell phone.)
- “What is the total of the percentages in the first column? What's wrong with that answer?” (205%. It doesn't make sense since it is greater than 100%. It is combining percentages of different wholes, so it doesn't make sense to add them.)
- “Did you say there was or was not an association between age and cell-phone use? Explain your reasoning.” (There is an association since the percentages for the 10 to 12 year olds is very different from the other two age groups.)
- “What did the relative frequency table show you that was harder to see in the original two-way table?” (It was much easier to see the association with the percentages than the numbers. For example, the number of people without a cell phone for the older 2 age groups was the same, but the percentages were a little different.)

Lesson Synthesis

Display the graphs and tell students to refer to the data from the tables in the Building Another Type of Two-Way Table activity.





Consider asking these discussion questions to emphasize the main ideas from the lesson:

- "Which graphical representation do you prefer for the data?" (It is easier to see the association with the segmented bar graph, but the actual values are lost.)
- "In the original table, what did the number 40 represent? How is that group of people represented in the other table and the two graphs?" (There are 40 people who are 13 to 15 years old that have a cell phone. In the relative frequency table, this is represented by the 80%. In the bar graph, this is represented by the taller, blue bar for set of bars labeled 13–15 years old. In the segmented bar graph, this is represented by the bottom, blue segment of the bar labeled 13–15 years old.)
- "What values in the tables represent the same information as the tallest, yellow-striped bar in the bar graph?" (The tallest yellow-striped bar represents the number of people who are 10–12 years old who do not have cell phones. In the original table, this is the 35 value and in the relative frequency table this is about 58%.)

9.4 Guitar and Golf

Cool Down: 5 minutes

Students create their own data based on how they view the students at their school. This data is then converted into a relative frequency table to practice the skill needed to create segmented bar graphs.

Addressing

- 8.SP.A.4

Launch

Tell students that they will make up their own numbers to fill in the two-way table so that the data reflects what they think is reasonable for students at your school. Prepare a two-way table to collect the raw data from the class that will be used in a future lesson.

Student Task Statement

1. In a class of 25 students, some students play a sport, some play a musical instrument, some do both, some do neither. Complete the two-way table to show data that might come from this class.

	plays an instrument	does not play an instrument	total
plays a sport			
does not play a sport	5		
total			25

2. Using the entries from the previous table, complete this table so that it shows relative frequencies. Round entries to the nearest percentage point.

	plays an instrument	does not play an instrument	total
plays a sport			
does not play a sport			

Student Response

Answers vary. Sample response:

1.

	plays an instrument	does not play an instrument	total
plays a sport	2	16	18
does not play a sport	5	2	7
total	7	18	25

2.

	plays an instrument	does not play an instrument	total
plays a sport	11% since $2 \div 18 \approx 0.11$	89% since $16 \div 18 \approx 0.89$	100%
does not play a sport	71% since $5 \div 7 \approx 0.71$	29% since $2 \div 7 \approx 0.29$	100%

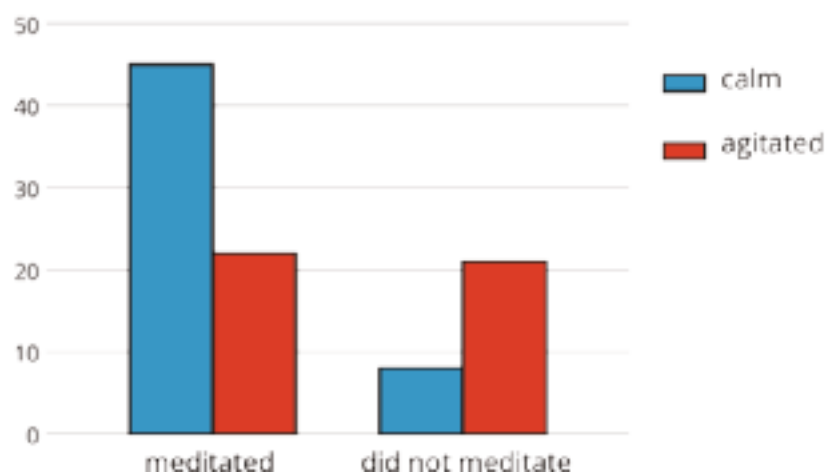
Student Lesson Summary

When we collect data by counting things in various categories, like red, blue, or yellow, we call the data *categorical data*, and we say that color is a *categorical variable*.

We can use **two-way tables** to investigate possible connections between two categorical variables. For example, this two-way table of frequencies shows the results of a study of meditation and state of mind of athletes before a track meet.

	meditated	did not meditate	total
calm	45	8	53
agitated	23	21	44
total	68	29	97

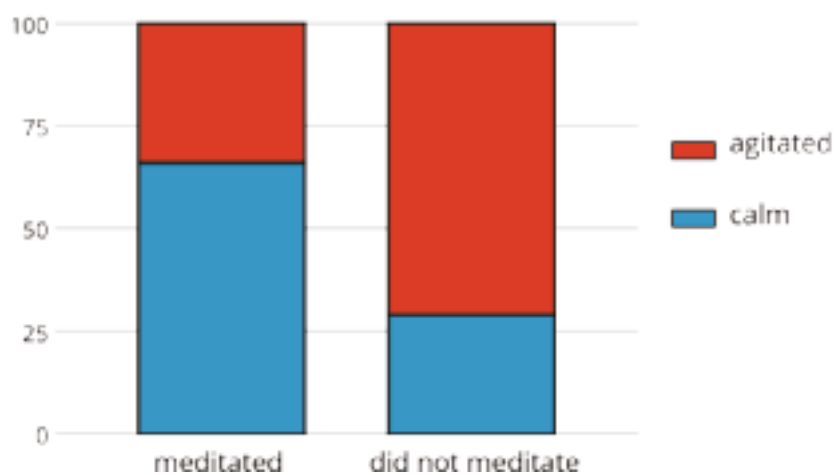
If we are interested in the question of whether there is an association between meditating and being calm, we might present the frequencies in a bar graph, grouping data about meditators and grouping data about non-meditators, so we can compare the numbers of calm and agitated athletes in each group.



Notice that the number of athletes who did not meditate is small compared to the number who meditated (29 as compared to 68, as shown in the table).

If we want to know the proportions of calm meditators and calm non-meditators, we can make a two-way table of **relative frequencies** and present the relative frequencies in a **segmented bar graph**.

	meditated	did not meditate
calm	66%	28%
agitated	34%	72%
total	100%	100%



Glossary

- relative frequency
- segmented bar graph
- two-way table

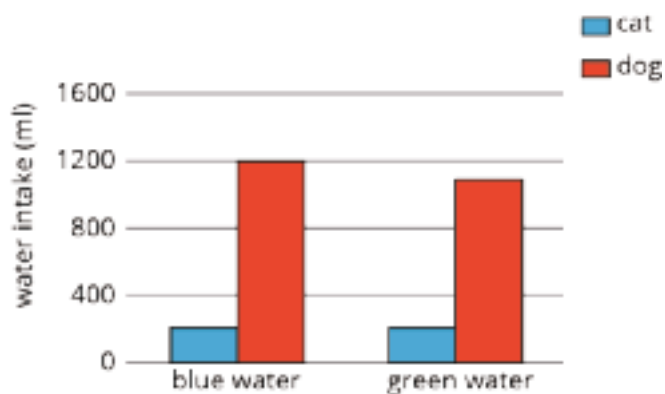
Lesson 9 Practice Problems

Problem 1

Statement

A scientist wants to know if the color of the water affects how much animals drink. The average amount of water each animal drinks was recorded in milliliters for a week and then graphed. Is there evidence to suggest an association between water color and animal?

	cat intake (ml)	dog intake (ml)	total (ml)
blue water	210	1200	1410
green water	200	1100	1300
total	410	2300	2710



Solution

No, the relative frequencies of the animals drinking each color of water are about the same, so there is no evidence of association.

Problem 2

Statement

A farmer brings his produce to the farmer's market and records whether people buy lettuce, apples, both, or something else.

	bought apples	did not buy apples
bought lettuce	14	58
did not buy lettuce	8	29

Make a table that shows the relative frequencies for each row. Use this table to decide if there is an association between buying lettuce and buying apples.

Solution

	bought apples	did not buy apples	total
bought lettuce	19% since $14 + 58 = 72$ and $14 \div 72 = 0.19\overline{4}$	81% since $58 \div 72 = 0.80\overline{5}$	100%
did not buy lettuce	22% since $8 + 29 = 37$ and $8 \div 37 = 0.21\overline{6}$	78% since $29 \div 37 = 0.78\overline{3}$	100%

Since the percentages in each column are close, there is not enough evidence to suggest an association between buying apples and lettuce.

Problem 3

Statement

Researchers at a media company want to study news-reading habits among different age groups. They tracked print and online subscription data and made a 2-way table.

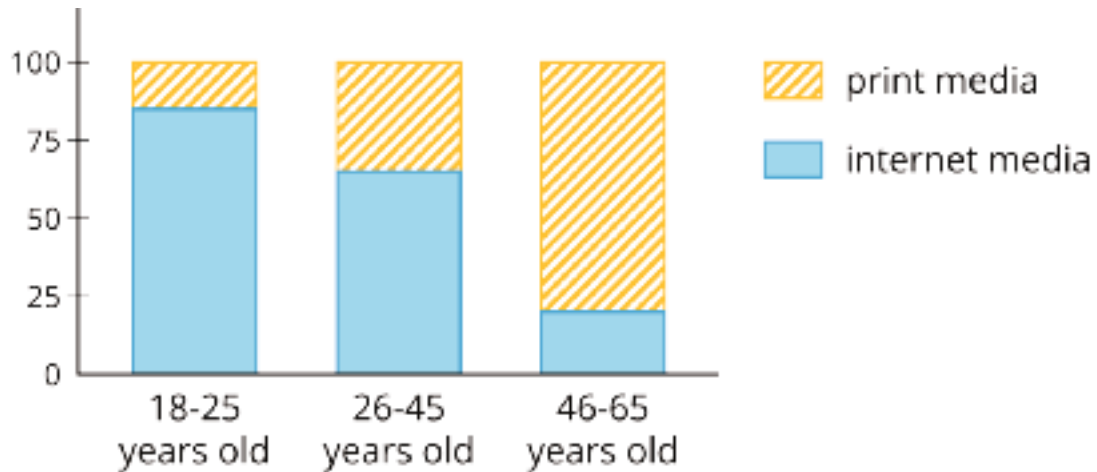
	internet media	print media
18-25 year olds	151	28
26-45 year olds	132	72
46-65 year olds	48	165

- a. Create a segmented bar graph using one bar for each row of the table.

- b. Is there an association between age groups and the method they use to read articles?
Explain your reasoning.

Solution

a.

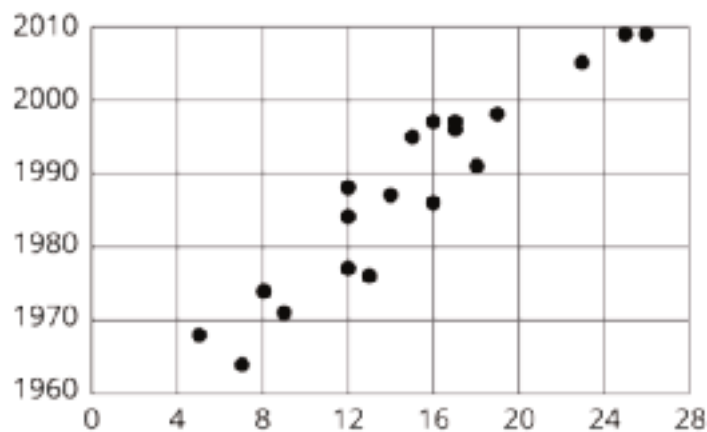


- b. Yes. Explanations vary. Sample explanation: The segments of the bars are not very close to being the same size. Younger age groups use the internet articles much more than print articles, while the oldest age group reverses that pattern.

Problem 4

Statement

Using the data in the scatter plot, what is a reasonable slope of a model that fits this data?



- A. -2.5
- B. -1
- C. 1
- D. 2.5

Solution

D

(From Unit 6, Lesson 6.)

Lesson 10: Using Data Displays to Find Associations

Goals

- Create a two-way table and a segmented bar graph that represent relative frequencies, and interpret (orally) the frequencies in context.
- Determine (in writing) whether categorical data has a positive, negative, or no association using a relative frequency table or segmented bar graph, and justify (orally) the reasoning.

Learning Targets

- I can create relative frequency tables, bar graphs, and segmented bar graphs from frequency tables to find associations among variables.

Lesson Narrative

In this lesson, students use two-way tables, bar graphs, and segmented bar graphs to decide whether there is evidence of an association in categorical data or not (MP4).

Alignments

Addressing

- 8.SP.A.4: Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?

Instructional Routines

- MLR1: Stronger and Clearer Each Time
- Think Pair Share

Required Materials

Colored pencils

Straightedges

A rigid edge that can be used for drawing line segments. Sometimes a ruler is okay to use as a

straightedge, but sometimes it is preferable to use an unruled straightedge, like a blank index card.

Required Preparation

Use the data from the previous lesson's cool-down to build a two-way table of students' responses. Provide access to materials for students to create their own segmented bar graphs including colored pencils and straightedges.

Student Learning Goals

Let's use data displays to find associations.

10.1 Sports and Musical Instruments

Warm Up: 5 minutes

The purpose of this warm-up is for students to answer questions about relative frequency of items after finding missing information in a two-way table.

Monitor for students who find the percentages for the final two questions using different strategies to share during the whole-class discussion.

Addressing

- 8.SP.A.4

Launch

Give students 2 minutes of quiet work time followed by a whole-class discussion.

Student Task Statement

For a survey, students in a class answered these questions:

- Do you play a sport?
- Do you play a musical instrument?

1. Here is a two-way table that gives some results from the survey. Complete the table, assuming that all students answered both questions.

	plays instrument	does not play instrument	total
plays sport	5		16
does not play sport			
total		15	25

2. To the nearest percentage point, what percentage of students who play a sport *don't* play a musical instrument?
3. To the nearest percentage point, what percentage of students who *don't* play a sport also *don't* play a musical instrument?

Student Response

1.		plays instrument	does not play instrument	total
	plays sport	5	11 since $16 - 5 = 11$.	16
	does not play sport	5 since $9 - 4 = 5$.	4 since $15 - 11 = 4$.	9 since $25 - 16 = 9$.
	total	10 since $25 - 15 = 10$.	15	25

2. 69% because $\frac{11}{16} = 0.6875$.

3. 44% because $\frac{4}{9} = 0.444\ldots$

Activity Synthesis

Ask students to share the missing information they found for the table. Record and display their responses for all to see.

Select students previously identified to explain how they found the percentages for the final two questions and what that percentage represents.

1. Students who find a percentage using the values given (for example 31% since $\frac{5}{16} \approx 0.31$), then subtract from 100% (for example 69% since $100 - 31 = 69$) to answer the question.
2. Students who find the actual values first by subtracting (for example $16 - 5 = 11$) then compute the percentage (for example 69% because $\frac{11}{16} = 0.6875$).

Ask the rest of the class if they agree or disagree with the strategies and give time for any questions they have.

10.2 Sports and Music Association

20 minutes

Now that students are more familiar with two-way tables showing relative frequency, they are ready to create their own segmented bar graphs. In this activity, students create two segmented bar graphs based on the same two-way table by considering percentages of the rows and columns separately. After creating the segmented bar graphs, they are analyzed to determine if there is an association present in the data.

Addressing

- 8.SP.A.4

Instructional Routines

- MLR1: Stronger and Clearer Each Time
- Think Pair Share

Launch

Arrange students in groups of 2. After a brief introduction, give 5–10 minutes of quiet work time. Ask students to compare their answers with their partner and try to resolve any differences. Finish with a whole-class discussion.

Display the two-way table from the previous lesson's cool-down activity containing the data collected about the class's playing sports and musical instruments. If the data is unavailable, the data from this lesson's warm-up can be used.

Tell students they should work with their partners to each work on one of the graphs. One student should work on problems 1 and 2 while their partner should work on 3 and 4. After they have completed their graphs, they should work together to understand their partners graphs and complete the last problem together.

Access for Students with Disabilities

Action and Expression: Internalize Executive Functions. Chunk this task into more manageable parts to support students who benefit from support with organization and problem solving. For example, present one question at a time. Some students may benefit from a checklist on how to create a segmented bar graph.

Supports accessibility for: Organization; Attention

Anticipated Misconceptions

Students may draw the segmented bar graph incorrectly. Most likely, they will accidentally graph frequency instead of relative frequency. They may also graph relative frequencies, but without stacking them. Both segmented bars should go from 0 to 100.

Student Task Statement

Your teacher will give you a two-way table with information about the number of people in your class who play sports or musical instruments.

1. Complete this table to make a two-way table for the data from earlier. The table will show relative frequencies *by row*.

	plays instrument	does not play instrument	row total
plays sport			100%
does not play sport			100%

2. Make a segmented bar graph for the table. Use one bar of the graph for each row of the table.



3. Complete the table to make a two-way table for the data from earlier. The table will show relative frequencies *by column*.

	plays instrument	does not play instrument
plays sport		
does not play sport		
column total	100%	100%

4. Using the values in the table, make a segmented bar graph. Use one bar of the graph for each column of the table.



5. Based on the two-way tables and segmented bar graphs, do you think there is an association between playing a sport and playing a musical instrument? Explain how you know.

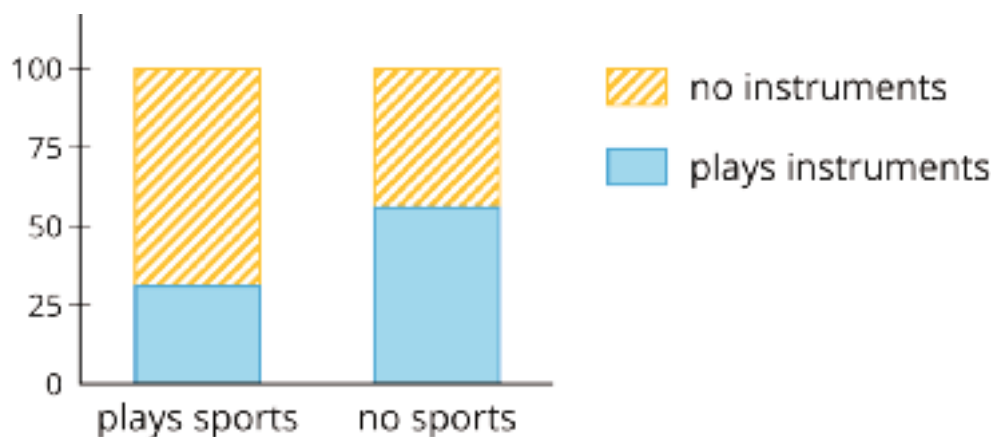
Student Response

Answers vary based on class data from previous cool-down. For the warm-up data:

1.

	plays instrument	does not play instrument	row total
plays sport	31%, since $6 \div 16 = 0.3125$	69%, since $11 \div 16 = 0.6875$	100%
does not play sport	56%, since $5 \div 9 = 0.\bar{5}$	44%, since $4 \div 9 = 0.\bar{4}$	100%

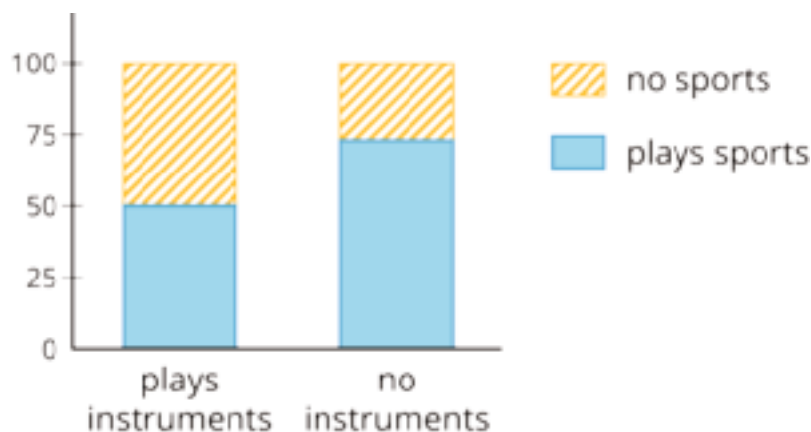
2.



3.

	plays instrument	does not play instrument
plays sport	50%, since $5 \div 10 = 0.5$.	73%, since $11 \div 15 = 0.7\bar{3}$.
does not play sport	50%, since $5 \div 10 = 0.5$.	27%, since $4 \div 15 = 0.2\bar{6}$.
column total	100%	100%

4.



5. I think there is a negative association between playing a sport and playing an instrument. Students who play a sport are less likely than their other classmates to play an instrument, and students who play an instrument are less likely than their other classmates to play a sport.

Activity Synthesis

To clarify how to create and interpret segmented bar graphs, ask:

- “What different information can be seen by the two segmented bar graphs?”
- “Why are the numbers in the top left box in the two tables different? What do they mean?” (In the first table it represents the percentage who also play musical instruments out of all the people who play sports. In the second table it represents the percentage of people who also play sports out of all the people who play musical instruments.)

- “Is there an association between the two variables? Explain or show your reasoning.” (The answer will depend on class data, but the reasoning should include an analysis of the relative frequencies within categories. There is an association if the percentages within one category are very different from the percentages in another category.)

If there is an association, ask what the segmented bar graphs would look like if there was no association. If there is not an association, ask what the segmented bar graphs would look like if there was one.

Access for English Language Learners

Writing, Speaking: MLR1 Stronger and Clearer Each Time. Use this routine to give students a structured opportunity to revise and refine 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 them strengthen their ideas and clarify their language (e.g., “Why do you think there is a (positive/negative) association?”, “How do the relative frequencies help to answer this question?”, “How could you say that another way?”, etc.). Students can borrow ideas and language from each partner to strengthen the final product. They can return to the first partner and revise and refine their initial response.

Design Principle(s): Optimize output (for explanation)

10.3 Colored Erasers

15 minutes

This activity provides students less structure for their work in creating segmented bar graphs to determine an association (MP4). In addition, the data in this activity is split into more than two options. Students work individually to create a segmented bar graph based on either columns or rows and then share their information with a partner who has created the other segmented bar graph. Together, partners discuss the segmented bar graphs to determine if there is an association between the variables (MP3). In particular, students should notice that there is evidence of an association is the relative frequencies within a category are very different from the relative frequencies in another category.

As students work, identify groups that use the different segmented bar graphs to explain why there is an association between the color of the eraser and flaws.

Addressing

- 8.SP.A.4

Instructional Routines

- Think Pair Share

Launch

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

Provide students access to colored pencils. Either assign or have partners choose which will make a graph for each row and which will make a graph for each column.

Access for Students with Disabilities

Representation: Access for Perception. Read the directions aloud. Students who both listen to and read the information will benefit from extra processing time. Check for understanding by inviting students to rephrase directions in their own words.

Supports accessibility for: Language

Student Task Statement

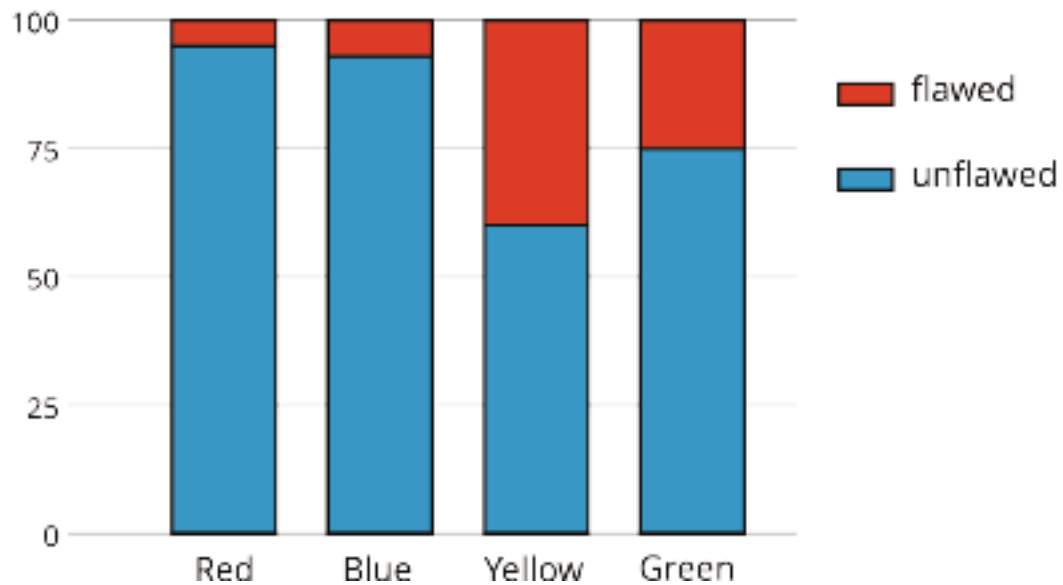
An eraser factory has five machines. One machine makes the eraser shapes. Then each shape goes through the red machine, blue machine, yellow machine, or green machine to have a side colored.

The manager notices that an uncolored side of some erasers is flawed at the end of the process and wants to know which machine needs to be fixed: the shape machine or some of the color machines. The manager collected data on the number of flawed and unflawed erasers of each color.

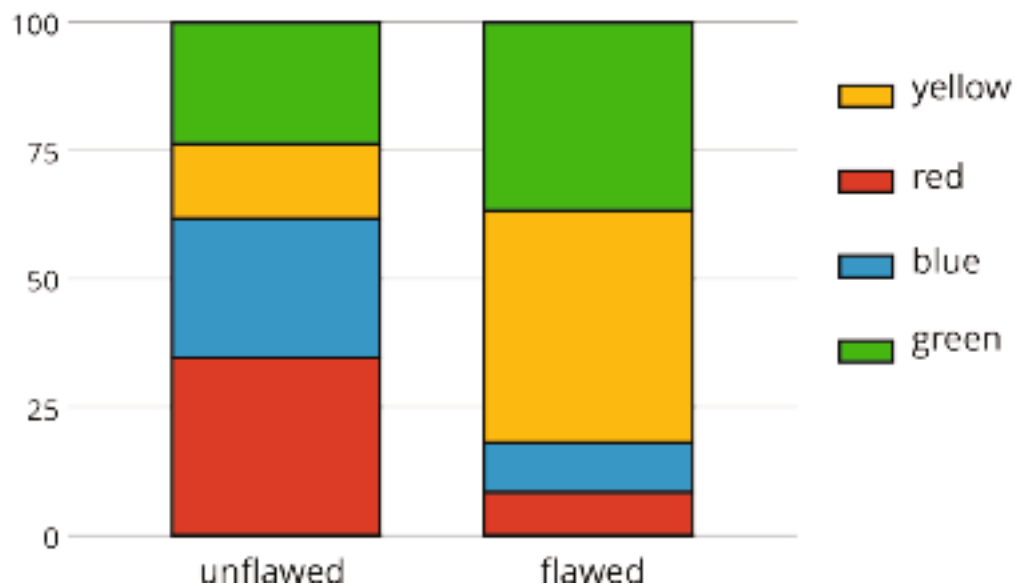
	unflawed	flawed	total
red	285	15	300
blue	223	17	240
yellow	120	80	200
green	195	65	260
total	823	177	1000

1. Work with a partner. Each of you should make one segmented bar graph for the data in the table. One segmented bar graph should have a bar for each *row* of the table. The other segmented bar graph should have one bar for each *column* of the table.
2. Are the flawed erasers associated with certain colors? If so, which colors? Explain your reasoning.

Student Response



1.



2. The flawed erasers are positively associated with the yellow and green erasers. A much higher percent of the total yellow and green erasers are flawed than red and blue are flawed. The yellow and green erasers also make up a larger percent of the total flawed erasers.

Are You Ready for More?

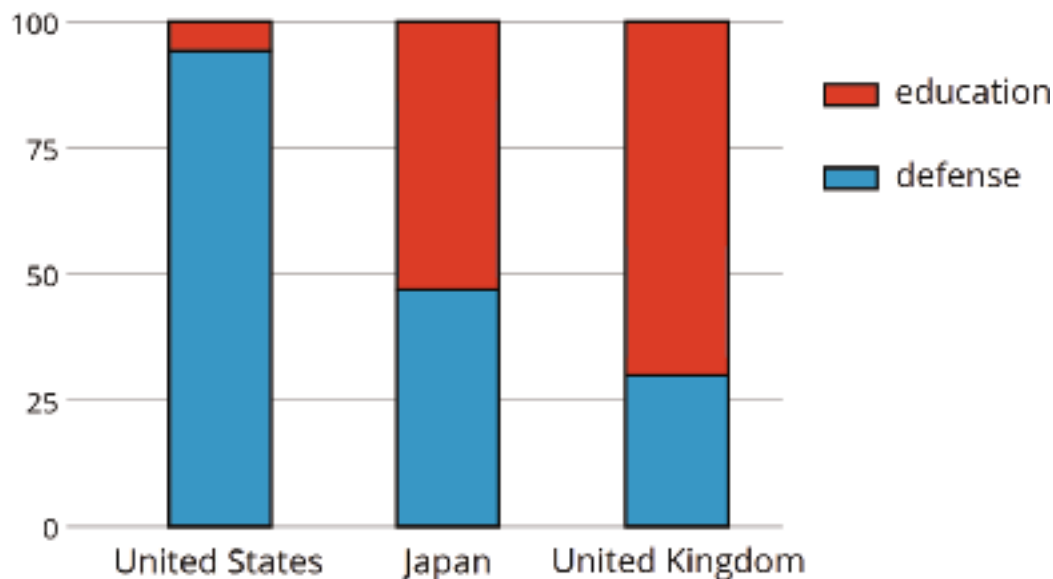
Based on the federal budgets for 2009, the table shows where some of the federal money was expected to go. The values are in billions of U.S. Dollars.

	United States	Japan	United Kingdom
defense	718.4	42.8	49.2
education	44.9	47.5	113.9

1. Why would a segmented bar graph be more useful than the table of data to see any associations between the country and where the money is spent?
2. Create a segmented bar graph that represents the data from the table.
3. Is there an association between the country's budget and their spending in these areas? Explain your reasoning.

Student Response

1. Answers vary. Sample response: The total amount of money spent by the United States is so much greater than the other countries that it would help to see the percentages rather than actual dollars spent.



- 2.
3. There is an association. The United States spends a much larger percentage of its budget on defense than on education relative to the other countries. Japan spends almost equal amounts on defense and education while the United Kingdom spends a lot more on education than on defense.

Activity Synthesis

The purpose of this discussion is to identify strategies for creating segmented bar graphs and for analyzing them to determine if there is an association among variables.

Ask, "What strategies did you use to create the segmented bar graphs?" (First, we created a new table of the relative frequencies. Then we approximated the heights of the segments based on the percentages from the table.)

Select previously identified groups to share their explanation for noticing an association.

1. Groups that use the segmented bar graph based on rows.
2. Groups that use the segmented bar graph based on columns.

After both explanations are shared, ask students, "Do you think that noticing the association was easier with one of the graphs?" (Likely the segmented bar graph based on rows is easier since there are only 2 segments and it is easier to see that the yellow and green erasers are more flawed.)

Finally, ask students, "If there was not an association between color and flaws, what might the segmented bar graph based on the rows look like? What might the segmented bar graph based on the columns look like?" (The segmented bar graph based on the rows would have each segmented bar look about the same. That is, the line dividing the two segments would be at about the same height in each bar. The segmented bar graph based on the columns would have segments that are all approximately equal. That is, each segment should represent about 25% of the entire bar.)

Lesson Synthesis

Remind students that we have been looking for associations in categorical data, and that there is evidence of an association if the relative frequencies of some characteristic are very different from each other in the different groups. Ask:

- "Is it easier to see evidence of an association in a frequency table or a *relative* frequency table?" (It depends on the data. If the two groups are approximately the same size, it doesn't matter very much, but when they are different sizes, it is usually easier to compare using relative frequencies.)
- "How can we see evidence of an association in a two-way table of either kind?" (By numerically comparing the proportions between the two groups.)
- "How can we see evidence of an association in a bar graph or segmented bar graph?" (By visually comparing the proportions between the two groups.)

10.4 Class Preferences

Cool Down: 5 minutes

Students are given a two-way table in which one variable has more than two outcomes and segmented bar graph to determine if there is an association between the two variables.

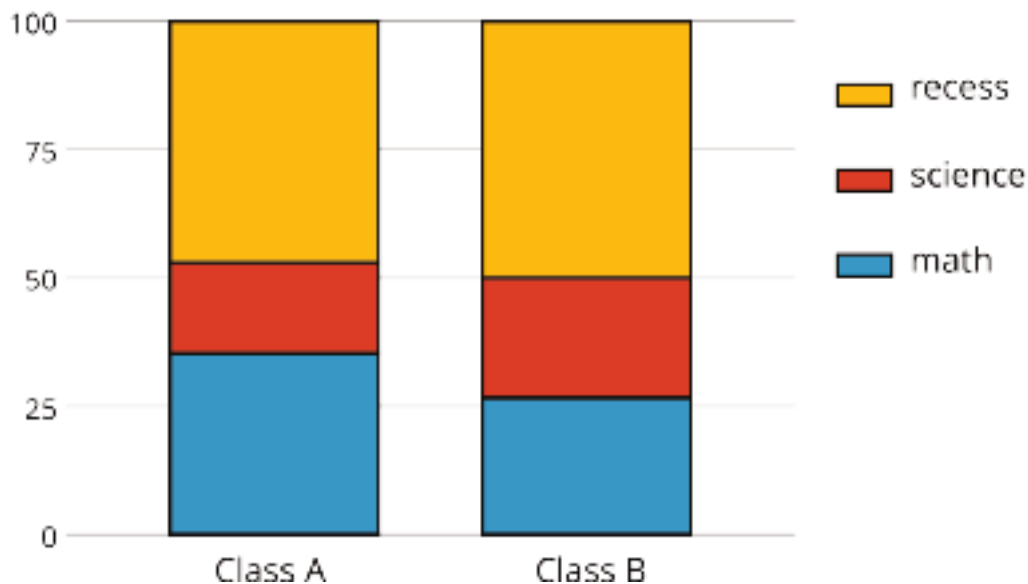
Addressing

- 8.SP.A.4

Student Task Statement

Here are a two-way table and segmented bar graph for data from students in two classes. Do they show evidence of differences between the two classes?

	prefers math	prefers science	prefers recess
class A	6	3	8
class B	8	7	15

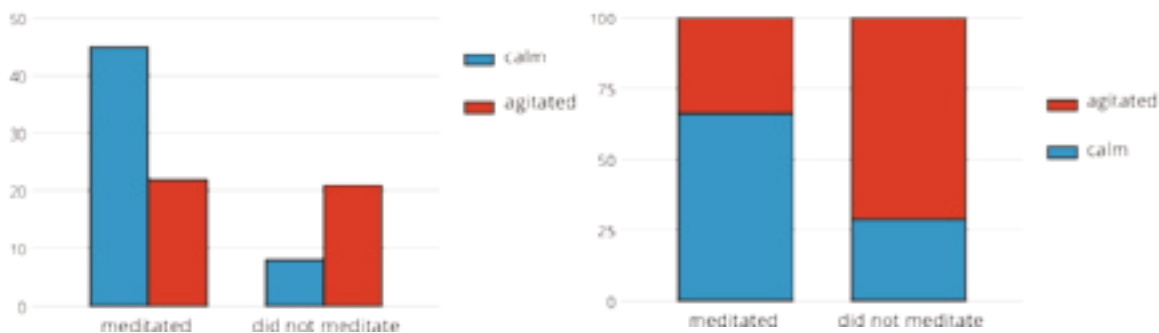


Student Response

There is no evidence of different preferences associated with each class because the segments in the bars are about the same size.

Student Lesson Summary

In an earlier lesson, we looked at data on meditation and state of mind in athletes.



Is there an association between meditation and state of mind?



The bar graph shows that more athletes were calm than agitated among the group that meditated, and more athletes were agitated than calm among the group that did not. We can see the proportions of calm meditators and calm non-meditators from the segmented bar graph, which shows that about 66% of athletes who meditated were calm, whereas only about 27% of those who did not meditate were calm.

This does not necessarily mean that meditation causes calm; it could be the other way around, that calm athletes are more inclined to meditate. But it does suggest that there is an association between meditating and calmness.

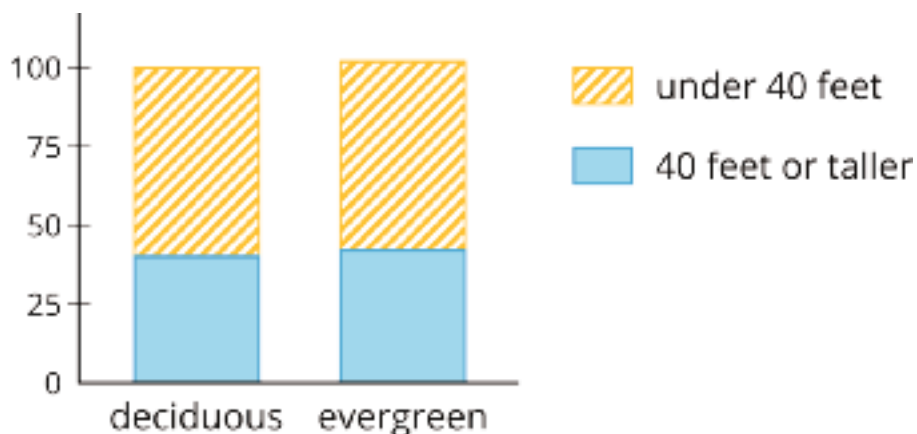
Lesson 10 Practice Problems

Problem 1

Statement

An ecologist is studying a forest with a mixture of tree types. Since the average tree height in the area is 40 feet, he measures the height of the tree against that. He also records the type of tree. The results are shown in the table and segmented bar graph. Is there evidence of an association between tree height and tree type? Explain your reasoning.

	under 40 feet	40 feet or taller	total
deciduous	45	30	75
evergreen	14	10	24
total	59	40	99



Solution

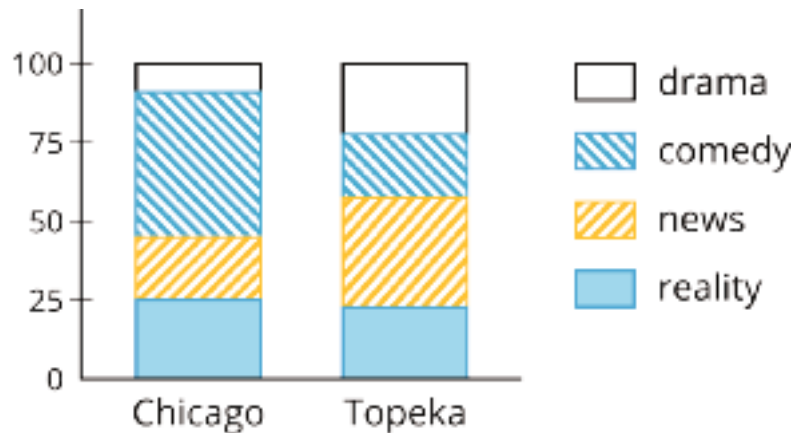
No, 60% of the deciduous trees are under 40 feet, and 40% are at least 40 feet. Similarly, 58% of evergreens are under 40 feet, and 42% are at least 40 feet. From the data recorded, there is not a clear association.

Problem 2

Statement

Workers at an advertising agency are interested in people's TV viewing habits. They take a survey of people in two cities to try to find patterns in the types of shows they watch. The results are recorded in a table and shown in a segmented bar graph. Is there evidence of different viewing habits? If so, explain.

	reality	news	comedy	drama
Chicago	50	40	90	20
Topeka	45	70	40	45



Solution

Yes, there are differences. Topekans watch news and dramas much more than Chicagoans. Chicagoans watch comedies much more. They watch about the same amount of reality TV.

Problem 3

Statement

A scientist is interested in whether certain species of butterflies like certain types of local flowers. The scientist captures butterflies in two zones with different flower types and records the number caught. Do these data show an association between butterfly type and zone? Explain your reasoning.

	zone 1	zone 2
eastern tiger swallowtail	16	34
monarch	24	46

Solution

No, there is not an association. 32% of eastern tiger swallowtails and 34% of monarchs were found in zone 1, so there is not a large difference in type of butterfly.

Section: Let's Put It to Work

Lesson 11: Gone In 30 Seconds

Goals

- Compare and contrast (orally) representations of bivariate data, including scatter plots, two-way tables, segmented bar graphs, and relative frequency tables.
- Describe (orally and in writing) associations in bivariate data using different representations of the same data.

Learning Targets

- I can collect data and analyze it for associations using scatter plots, two-way tables, and segmented bar graphs.

Lesson Narrative

In this optional culminating lesson, students collect, represent, and analyze quantitative bivariate data. Then they group the quantitative data to create two categories, and look for evidence of an association in the resulting categorical data. This lesson allows students the opportunity to carryout the entire process of data analysis beginning with data collections and ending with an analysis of graphical representations to determine whether there is an association between two variables in the data. Students attempt to describe their reaction times by using the various techniques from the unit to model (MP4) their data.

Alignments

Addressing

- 8.SP.A: Investigate patterns of association in bivariate data.

Instructional Routines

- MLR8: Discussion Supports

Required Materials

Stopwatches

Required Preparation

Students will work in groups of 4. Each group will need one stopwatch.

Student Learning Goals

Let's gather and analyze some timing data.

11.1 Measuring 30 Seconds

Optional: 45 minutes

To close the unit, students practice using the methods they have learned. Students begin by collecting data in small groups and then analyzing the data for the entire class. They create a scatter plot and two-way tables to look for patterns in the data (MP2). Although it is expected there is a very weak, if any, association in the data, the discussion asks students to think about what it would mean if there were a positive or negative association in context.

Addressing

- 8.SP.A

Instructional Routines

- MLR8: Discussion Supports

Launch

Arrange students in groups of 4. Select one student from each group to submit the data from the group when they finish recording their data. Prepare a table for all to see that is large enough to include the data from all of the students in the class.

Tell students that the same data is treated as both numerical and categorical data at different times within this activity. This is sometimes done when there is a threshold that makes sense for the context. In this case, it makes sense to look at data for which the second time is greater or less than the first time. Although the actual values may be lost when viewing the data as categorical, it can provide a different type of insight. In this example, we can more easily see if there was an improvement in reaction time for the second attempt.

Access for Students with Disabilities

Engagement: Internalize Self Regulation. Provide a project checklist that chunks the various steps of the project into a set of manageable tasks.

Supports accessibility for: Organization; Attention

Access for English Language Learners

Representing, Conversing: MLR8 Discussion Supports. Demonstrate the process for one round of the experiment, in which a group of students and staff play an example round while the rest of the class observes. This will help invite more student participation, conversations, and meta-awareness of language involving collecting and analyzing the data.

Design Principle(s): Support sense-making; Maximize meta-awareness.

Student Task Statement

In this activity, you'll get two chances to guess at how long 30 seconds is, then look for an association between the two guesses of all students.

1. Work with a partner. Follow the instructions listed here to gather your data.
 - One of you will hold a stopwatch where the other person cannot see it.
 - The person holding the stopwatch says "go" and starts the timer.
 - The other person says "stop" when they think 30 seconds have passed.
 - The person holding the stopwatch will stop the timer, then report and record the time to the nearest second.
 - The person holding the stopwatch will give a second chance, repeating the experiment.
 - After *both* times are recorded, switch roles.
2. Record the group data in this table. When you finish, a group member should give the data to the teacher.

name	time 1	time 2

3. Look at your data. Comparing Time 1 to Time 2, do you think there is a positive association, a negative association, or no association? Discuss your thinking with your group.
4. What are some ways you could organize and represent the entire class's data?

5. Make a scatter plot of the entire class's data and look for patterns. Identify any outliers and the type of any association you observe.



6. Draw two lines on your scatter plot: a vertical line and a horizontal line, each representing 30 seconds for one trial. Use the table for the class's data to complete this two-way table.

	time 2 < 30 sec	time 2 = 30 sec	time 2 > 30 sec	total
time 1 < 30 sec				
time 1 = 30 sec				
time 1 > 30 sec				
total				

7. Use the two-way table to decide whether there is an association between Time 1 and Time 2. Explain how you know.

Student Response

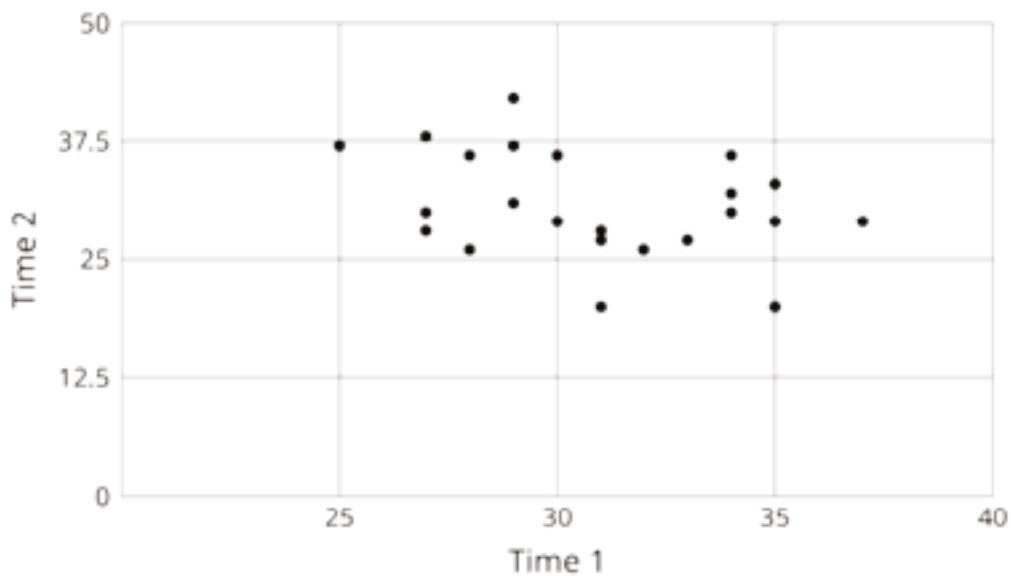
Answers vary based on student data. Sample response:

1. No response necessary.

2.

name	time 1	time 2
Jada	28	36
Lin	37	29
Diego	29	37
Andre	35	29

3. There appears to be a negative association. People who guessed too high the first time may have tried to stop sooner the second time, and people who guessed too low the first time may have waited longer to stop the second time.
4. We could show the class data as a scatter plot to show both times at once. We could find the difference of each time from exactly 30 seconds and show that with separate histograms, dot plots, or box plots.



5. There appears to be no association in the class data. There aren't any outliers, because all of the data was in a cluster around the point (30, 30).
- 6.

	time 2 < 30 sec	time 2 = 30 sec	time 2 > 30 sec
time 1 < 30 sec	2	1	9
time 1 = 30 sec	1	0	1
time 1 > 30 sec	2	1	9

7. There is not evidence of association, because the relative frequencies of one time greater and less than 30 seconds is about the same regardless of what the other time was.

Activity Synthesis

The purpose of the discussion is to solidify understanding of the strategies and tools from the unit. Specifically, how to organize and analyze categorical data to determine if there is an association between two variables.

Some questions for discussion:

- "What representation (data table, scatter plot, two-way table, or something else) helped the most in determining any relationships between time 1 and time 2?"
- "If someone's time was slower on the second attempt, how would that appear on the scatter plot?" (The point would be above and to the left of the line $y = x$, since the value for the y -coordinate would be greater than the value for the x -coordinate like the point (1.2, 1.4).)
- "If there was a positive association between the variables, what would that mean?" (In general, the slower the first time, the slower the second time. The faster the first time, the faster the second time as well.)
- "If there was a negative association between the variables, what would that mean?" (In general, the slower the first time, the faster the second time. The faster the first time, the slower the second time.)
- "If the class showed a tendency to overcompensate by going from a time that was longer than 30 seconds to one that was shorter or the other way around, how would that appear in the two-way table?" (The numbers in the table would be greatest on the bottom left and top right of the table with smaller numbers elsewhere.)

Lesson Synthesis

The goal of this discussion is to help students reflect on collecting, representing, and analyzing data with two linked variables. Display the data from the class for all to see while discussing these questions:

- “How is the data used in this unit different from data that would be displayed on a dot plot?” (In this unit we studied two variables for the same individuals while a dot plot would only show 1 variable.)
- “What are some ways we have graphically represent the data from this unit? Why might we choose one representation instead of another?” (We used tables and scatter plots with numerical data, like height and hand span. We used two-way tables and bar graphs with categorical data, like if someone plays an instrument or has a cell phone.)
- “How do we look for associations between variables?” (We look at trends in scatter plots when the data are numerical. We look for different relative frequencies when the data are categorical.)



Family Support Materials

Family Support Materials

Associations in Data

Here are the video lesson summaries for Grade 8, Unit 6: Associations in Data. Each video highlights key concepts and vocabulary that students learn across one or more lessons in the unit. The content of these video lesson summaries is based on the written Lesson Summaries found at the end of lessons in the curriculum. The goal of these videos is to support students in reviewing and checking their understanding of important concepts and vocabulary. Here are some possible ways families can use these videos:

- Keep informed on concepts and vocabulary students are learning about in class.
- Watch with their student and pause at key points to predict what comes next or think up other examples of vocabulary terms (the bolded words).
- Consider following the Connecting to Other Units links to review the math concepts that led up to this unit or to preview where the concepts in this unit lead to in future units.

Grade 8, Unit 6: Associations in Data	Vimeo	YouTube
Video 1: Using Scatter Plots to Visualize Data (Lessons 1–3)	Link	Link
Video 2: Using Lines to Model Data (Lesson 4–8)	Link	Link
Video 3: Associations in Categorical Data (Lessons 9–10)	Link	Link

Video 1

Video 'VLS G8U6V1 Using scatter plots to visualize data (Lessons 1–3)' available here:
<https://player.vimeo.com/video/500190466>.

Video 2

Video 'VLS G8U6V2 Using Lines to Model Data (Lesson 4–8)' available here:
<https://player.vimeo.com/video/502223668>.

Video 3

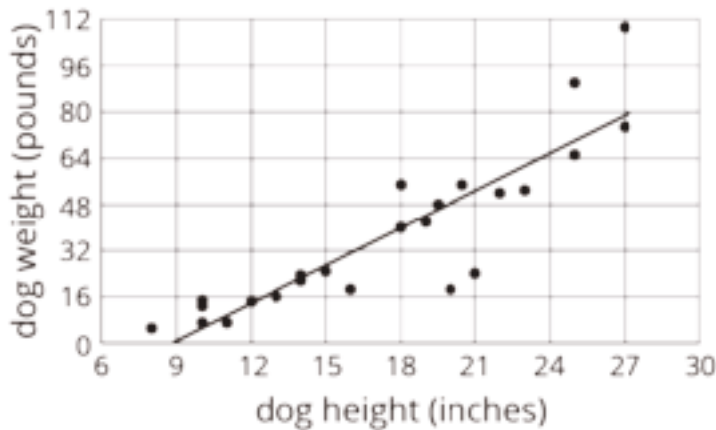
Video 'VLS G8U6V3 Associations in Categorical Data (Lessons 9–10)' available here:
<https://player.vimeo.com/video/507557063>.

- *Coming soon*

Does This Predict That?

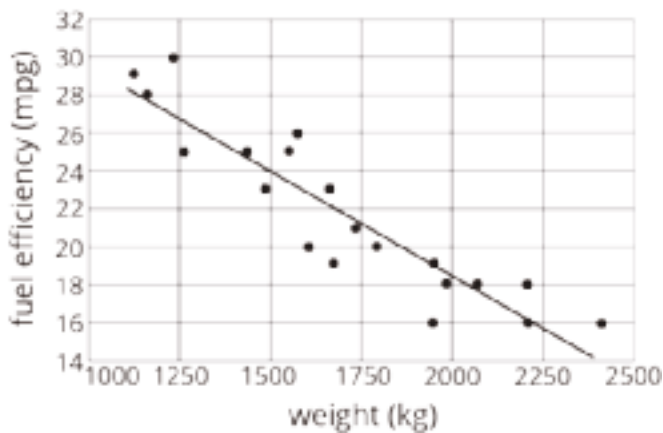
Family Support Materials 1

This week your student will work with **scatter plots**. Scatter plots show us how two different variables are related. In the example below, each plotted point corresponds to a dog, and its coordinates tell us the height and weight of that dog. The point on the lower left of the graph, for example, might represent a dog that is 8 inches tall and weighs about 5 pounds. The plot shows that, generally speaking, taller dogs weigh more than shorter dogs.



Since a larger value for one characteristic (height) generally means a larger value for the other characteristic (weight), we say that there is a **positive association** between dog height and dog weight.

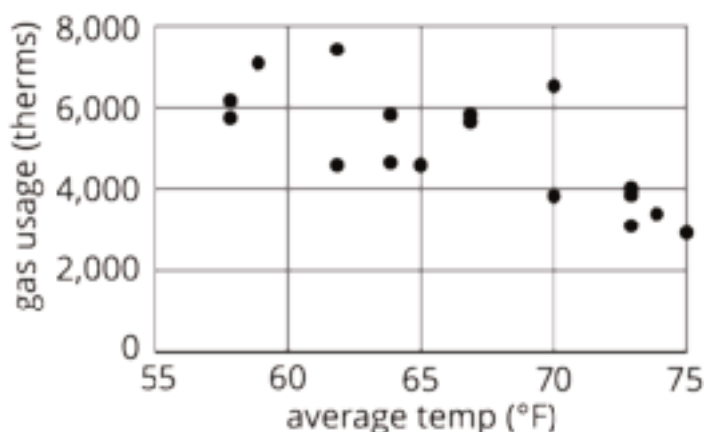
In the next example, each point corresponds to a car, and its coordinates tell us the weight and fuel efficiency of the car.



This time, we see that larger values for one characteristic (car weight) generally have lower values for the other characteristic (fuel efficiency), and so we say that there is a **negative association** between car weight and fuel efficiency.

Here is a task to try with your student:

The following scatter plot shows the relationship between average temperature and gas usage in a building.



1. How many points in the graph describe the building on 70-degree days?
Approximately how much gas was used on each of these days?
2. Do the variables in the gas usage for a building scatter plot show a positive association or a negative association?
3. On a 78-degree day, would the building be most likely to use (a) 1,800 therms of gas, (b) 4,200 therms of gas, or (c) 5,800 terms of gas?

Solution:

1. There are two points that describe gas usage for 70-degree days. On one of those days, the building used a little less than 4,000 therms of gas. On the other, the building used a little more than 6,000 therms.
2. Since less gas is used on warmer days, there is a negative association.
3. Following the trend in the graph, the building would likely use about 1,800 therms on a 78-degree day. You may draw in a line as in the dog and car scatter plots to help see this.

Associations in Categorical Data

Family Support Materials 2

This week your student will use two-way tables. Two-way tables are a way of comparing two variables. For example, this table shows the results of a study of the relation between meditation and state of mind of athletes before a track meet.

	meditated	did not meditate	total
calm	45	8	53
agitated	23	21	44
total	68	29	97

23 of the people who meditated were agitated, while 21 of the people who did not meditate were agitated. Does this mean that meditation has no impact or even a slight negative association with mood? Probably not. When we look for associations between variables it can be more informative to know the percentages in each category, like this:

	meditated	did not meditate
calm	66%	28%
agitated	34%	72%
total	100%	100%

Of the people who meditated, 66% were calm, and 34% were agitated. When we compare that to the percentages for people who did not meditate, we can now see more easily that the group of people who meditated has a lower percentage of athletes who are agitated. The percentages in this table are called **relative frequencies**.

Here is a task to try with your student:

The following table contains data about whether people in various age groups use their cell phone as their main alarm clock.

	use cell phone as alarm	do not use cell phone as alarm	total
18 to 29 years old	47	16	63
30 to 49 years old	66	23	87
50+ years old	31	39	70
total	144	78	220

1. Fill in the blanks in the table below with the relative frequencies for each row. These will tell us the percentage of people in each age group who use their phone as an alarm.

	use cell phone as alarm	do not use cell phone as alarm	total
18 to 29 years old	75%, since $\frac{47}{63} = 0.75$		100%
30 to 49 years old			
50+ years old			

2. Comparing just the 18 to 29 year olds and the 30 to 49 year olds, is there an association between cell phone alarm use and age?
3. Comparing the two youngest age brackets with the 50+ age bracket, is there an association between cell phone alarm use and age?

Solution:

1.

	use cell phone as alarm	do not use cell phone as alarm	total
18 to 29 years old	75%, since $\frac{47}{63} = 0.75$	25%, since $\frac{16}{63} = 0.25$	100%
30 to 49 years old	76%, since $\frac{66}{87} = 0.76$	24%, since $\frac{23}{87} = 0.24$	100%
50+ years old	44%, since $\frac{31}{70} = 0.44$	56%, since $\frac{39}{70} = 0.56$	100%

2. No: the relative frequencies are very similar.

3. Yes: using a cell phone as an alarm is associated with being in the younger age brackets. About 75% of 18 to 29 and 30 to 49-year olds use their cell phone as an alarm, but only 44% of people 50 years or older do.

Unit Assessments

Check Your Readiness A and B
End-of-Unit Assessment A and B

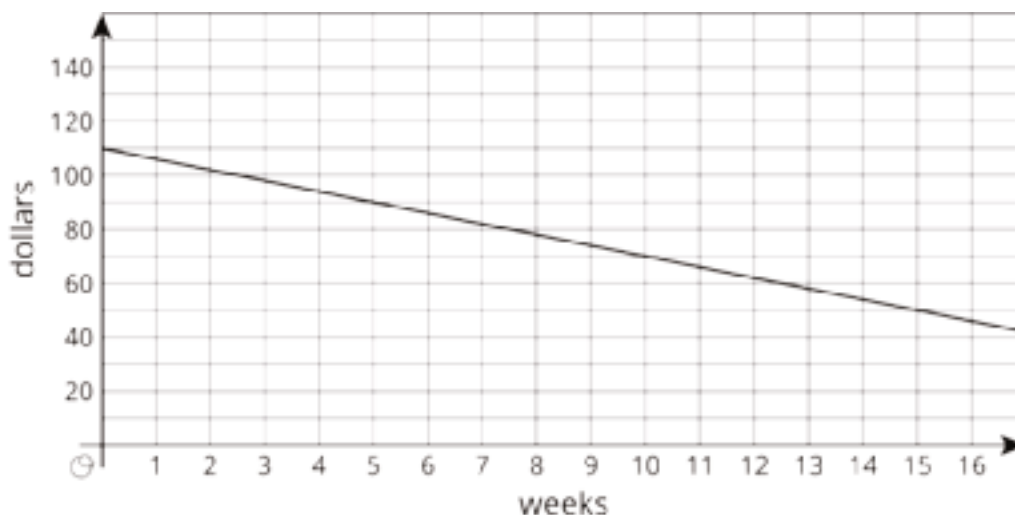
Associations in Data: Check Your Readiness (A)

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

1. A line contains the points $(-3, -2)$ and $(7, 2)$.
 - a. Without doing any calculations, determine whether the slope of this line is positive or negative. Explain or show your reasoning.
 - b. Calculate the slope of the line.
2. Diego has some money in his bank account before he starts a summer job. He deposits the money from the summer job in his bank account and doesn't spend any of it. After working 3 hours, he has \$71 in the account. After working 12 hours, he has \$134 in the account.

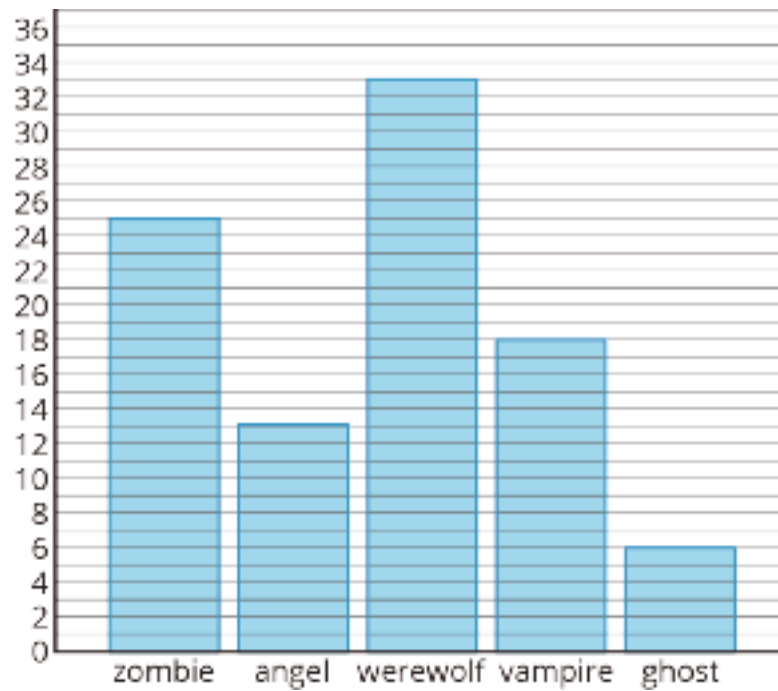
How much money does Diego earn per hour?

3. Here is a graph showing the balance in someone's savings account since the beginning of the year.



- a. Write an equation for the line shown on the graph.
- b. What does the slope mean in the situation?
- c. What does the vertical intercept mean in the situation?
4. In many schools, students have the choice between taking art, music, or some other elective.
- a. At Euclid Middle School, there are 200 students in the 8th grade. 40 students are taking art. What percentage of 8th graders at Euclid Middle School are taking art? Explain or show your reasoning.
- b. At Newton Middle School, there are 320 students in the 8th grade. 54 are taking music. What percentage of 8th graders at Newton Middle School are taking music? Explain or show your reasoning.

5. Students voted for their favorite entry in a Halloween costume contest. Jada recorded the results in a bar graph:



- Who got more votes: the angel or the zombie?
- How many votes did the vampire get?
- Who won the contest?

6. Students at Elena's school were polled about the animal they would most like to have as a pet.

animal	votes
bird	22
cat	45
dog	55
fish	37
rabbit	15

Make a bar graph that displays this information.

Associations in Data: Check Your Readiness (B)

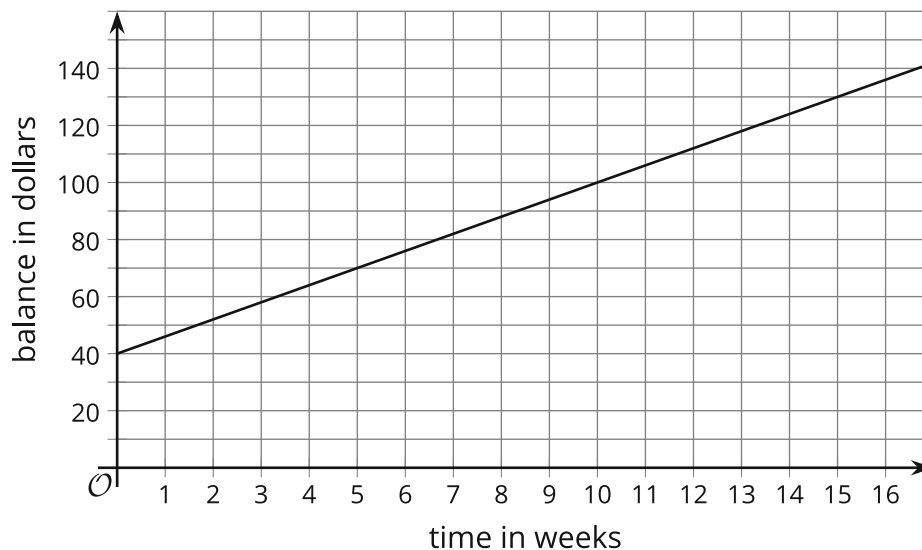
You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

1. A line contains the points $(-1, 4)$ and $(5, -3)$.
 - a. Without doing any calculations, determine whether the slope of this line is positive or negative. Explain or show your reasoning.

 - b. Calculate the slope of the line.

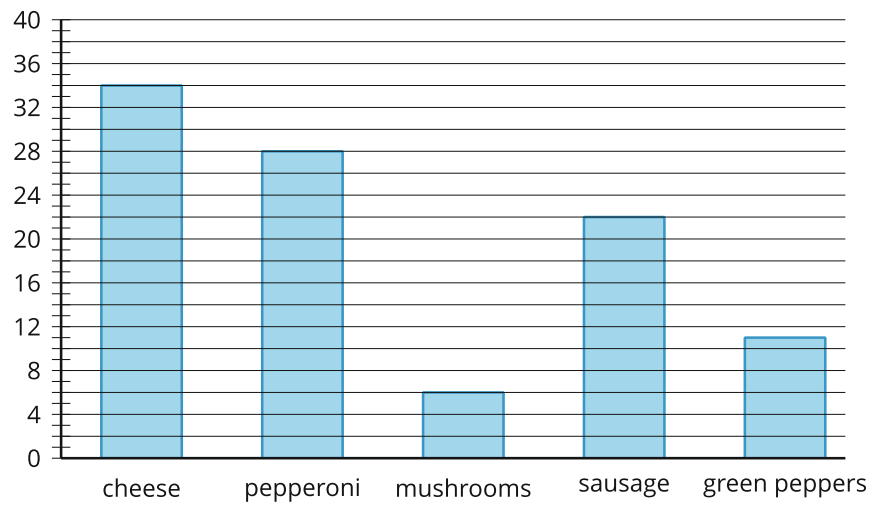
2. Lin has some money in her bank account before she starts a summer job. She deposits the money from the summer job in her bank account and doesn't spend any of it. After working 4 hours, she has \$85 in the account. After working 15 hours, she has \$217 in the account. How much money does Lin earn per hour?

3. Here is a graph showing the balance in someone's savings account since the beginning of the year.



- Write an equation for the line shown on the graph.
 - What does the slope mean in the situation?
 - What does the vertical intercept mean in the situation?
4. Determine each answer to the nearest tenth of a percent.
- 8 out of 32 football teams are on the East Coast. What percentage of football teams are on the East Coast?
 - It rained 4 out of 7 days this week. What percentage of days did it rain?

5. Students voted for their favorite pizza toppings. Kiran recorded the results in a bar graph.



- Which topping won the contest?
- How many votes did pepperoni get?
- Which topping got more votes: mushrooms or green peppers?

6. Students at Andre's school were polled about the mascot they would most like to have for their school.

mascot	number of votes
badger	14
hornet	25
viking	67
cardinal	49
wildcat	83

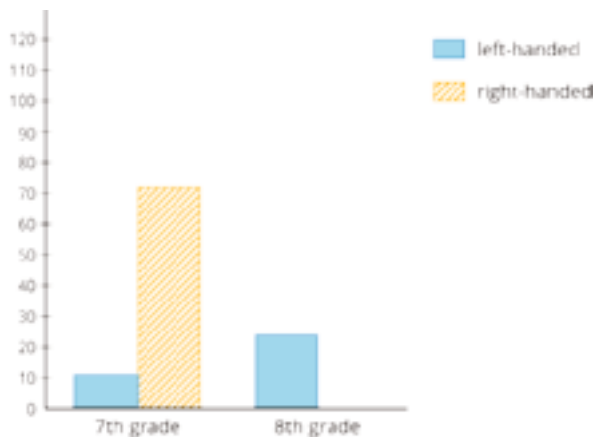
Make a bar graph that displays this information.

Associations in Data: End-of-Unit Assessment (A)

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

1. Noah gathered data at his school among 7th and 8th graders to see if there was an association between grade level and handedness. The table and graph show his data, but the number of right-handed 8th graders is missing.

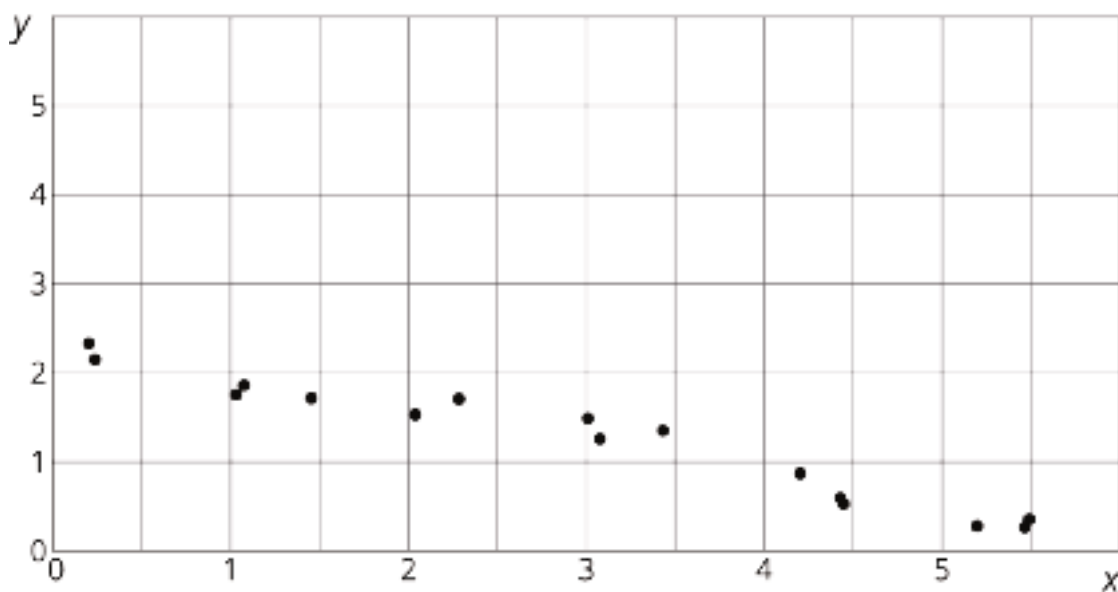
	left-handed	right-handed
7th grade	11	72
8th grade	24	



Noah found there was *no evidence* of an association between grade level and handedness. Which of these could be the number of right-handed 8th graders?

- A. 33
- B. 85
- C. 107
- D. 157

2. Here is a scatter plot:



The graph of what linear equation is a good fit for this data?

A. $y = -\frac{1}{3}x + 2$

B. $y = -\frac{1}{3}x + 6$

C. $y = \frac{1}{3}x + 2$

D. $y = \frac{1}{3}x + 6$

3. Select **all** the relationships that demonstrate a positive association between variables.

- A. Outside temperature and cost to heat a home
- B. Length of time you have walked and distance you have traveled
- C. Pounds of cherries you buy and amount of money you spend on cherries
- D. Speed of a train and the amount of time it takes for the train to get to its destination
- E. Number of people in a grocery check-out line and how long you have to wait to check out

4. a. Draw a scatter plot that shows a positive, linear association and has one clear outlier. Circle the outlier.

b. Draw a scatter plot that shows a negative association that is not linear.

5. Jada surveyed all 7th and 8th graders at her school about whether they have pets. Complete the missing entries in this two-way table.

	has pet	has no pet	total
7th grade	102		150
8th grade		68	175
total			

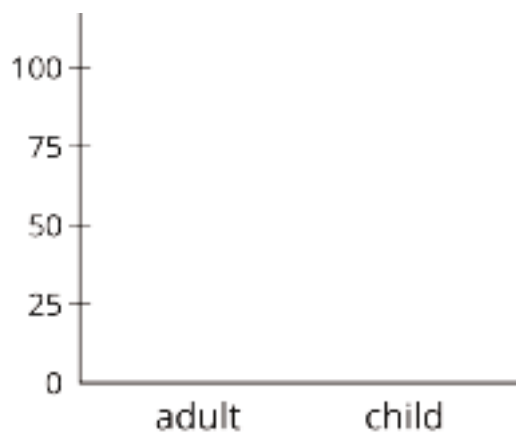
6. At a school social, children attend with family members. Everyone had a choice between a sweet snack and a salty snack. Here is a two-way table showing the number of adults and children who made each choice of snack.

	sweet snack	salty snack	total
adult	57	88	145
child	77	31	108
total	134	119	253

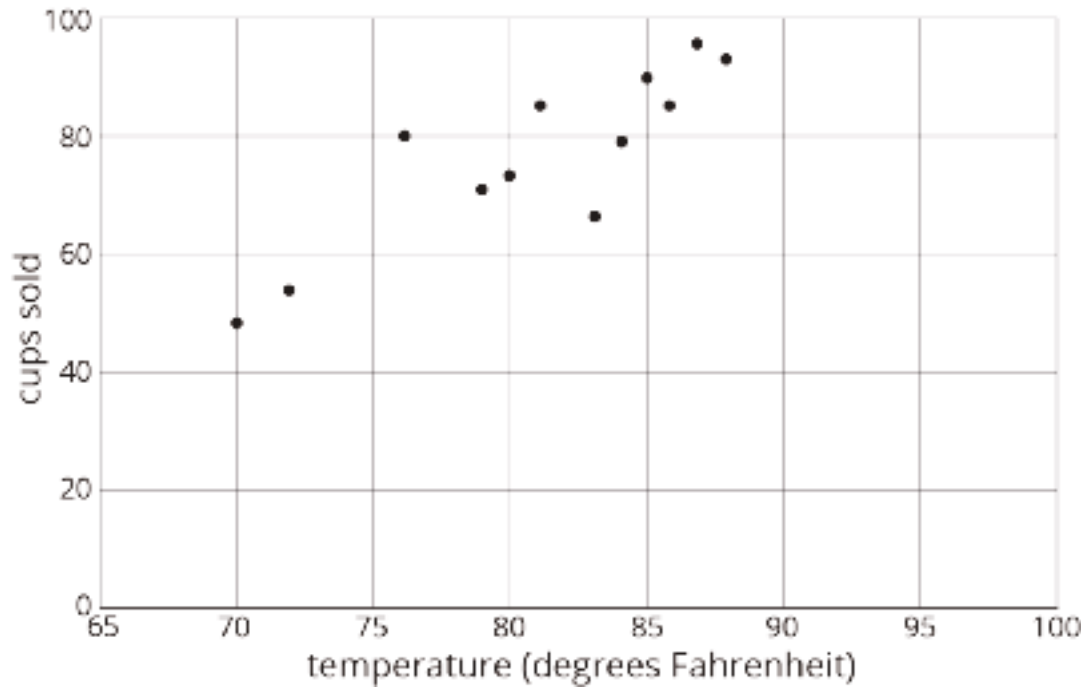
- a. Complete the table with relative frequency by row. Round to the nearest percent.

	sweet snack	salty snack	total
adult			100%
child			100%

- b. Make a segmented bar graph to represent the data in your table. Use one bar for each row of the table.



7. Lin opened a lemonade stand during the summer. She noticed that she sold more lemonade on warmer days. For each day she sold lemonade, she plotted the point (t, c) , where t represents high temperature and c represents cups of lemonade sold.



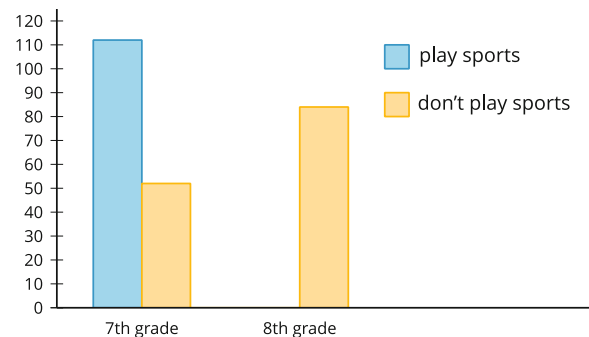
- On the same axes, draw a line that you think is a good fit for the data.
- A computer program found that the line $c = 2t - 89$ is a good fit for the data. Use this equation to predict how many cups of lemonade Lin might sell on a day when the high temperature is 74 degrees.
- The high temperature this Sunday is expected to be 5 degrees warmer than the high temperature this Saturday. Using the line $c = 2t - 89$, how many more cups of lemonade should Lin expect to sell on Sunday than Saturday? Explain or show your reasoning.

Associations in Data: End-of-Unit Assessment (B)

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

1. Jada gathered data at her school among 7th and 8th graders to see if there was an association between grade level and playing sports. The table shows her data, but the number of 8th graders who play sports is missing.

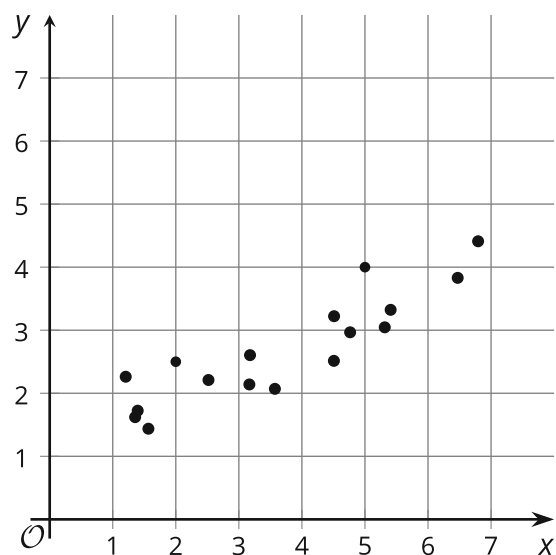
	play sports	don't play sports
7th graders	112	52
8th graders		84



Jada found there was *no evidence* of an association between grade level and playing sports. Which of these could be the number of 8th graders who play sports?

- A. 248
- B. 181
- C. 144
- D. 69

2. Here is a scatter plot:



The graph of what linear equation is a good fit for this data?

A. $y = \frac{1}{2}x$

B. $y = -\frac{1}{2}x + 1$

C. $y = -\frac{1}{2}x$

D. $y = \frac{1}{2}x + 1$

3. Select **all** the relationships that demonstrate a negative association between variables.

A. Number of absences from school and final grades

B. Outside temperature and ice cream sales

C. Price of houses and house sales

D. Number of rainy days and car accidents

E. Number of hours playing video games and grades

4. a. Draw a scatter plot that shows a negative, linear association and has one clear outlier. Circle the outlier.

b. Draw a scatter plot that shows a positive association that is not linear.

5. Andre surveyed all 7th and 8th grade students at his school about whether they bring their lunch or eat the school lunch. Complete the missing entries in this two-way table.

	bring home lunch	eat school lunch	total
7th grade students		219	293
8th grade students	168		261
total			

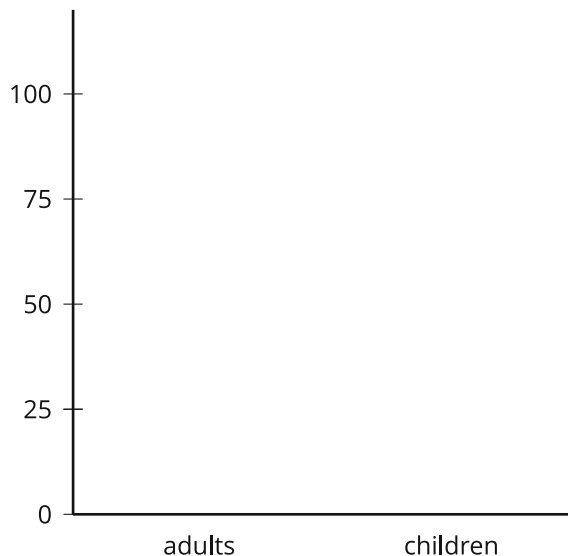
6. At the school carnival, adults attended with children. Everyone had a choice between a piece of pizza and a hot dog. Here is a two-way table showing the number of adults and children who made each choice.

	pizza	hot dog	total
adult	48	105	153
child	92	73	165
total	140	178	318

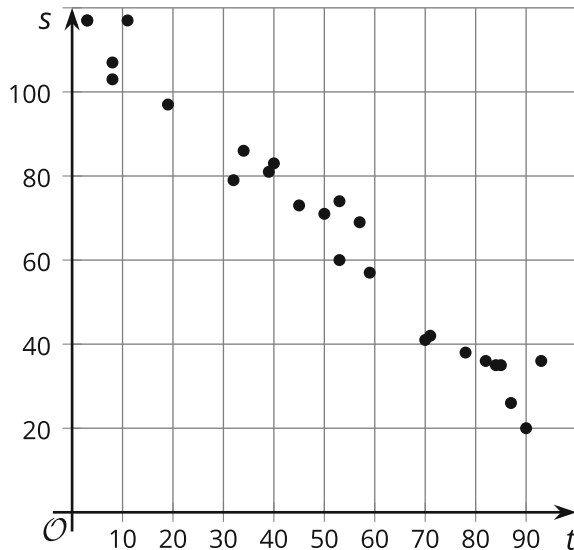
- a. Complete the table with relative frequency by row. Round to the nearest percent.

	pizza	hot dog	total
adult			100%
child			100%

- b. Make a segmented bar graph to represent the data in your table. Use one bar for each row of the table.



7. A deli owner noticed that as the outside temperature increased, they sold less soup. For each day soup was sold, they plotted the point (t, s) , where t represents high temperature and s represents bowls of soup sold.



- On the same axes, draw a line that you think is a good fit for the data.
- The deli owner found that the line $s = -t + 120$ is a good fit for the data. Use this equation to predict how many bowls of soup they might sell on a day when the high temperature is 32 degrees.
- The high temperature this Saturday is expected to be 10 degrees colder than the high temperature this Friday. Using the line $s = -t + 120$, how many more bowls of soup should the deli expect to sell on Saturday than Friday? Explain or show your reasoning.

Assessment Answer Keys

Check Your Readiness A and B
End-of-Unit Assessment A and B

Assessment Answer Keys

Assessment : Check Your Readiness (A)

Teacher Instructions

Graphing calculators should not be used. Use of a four-function or scientific calculator is acceptable. Provide access to graph paper.

Student Instructions

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

Problem 1

The content assessed in this problem is first encountered in Lesson 5: Describing Trends in Scatter Plots.

In this unit, students will need to eyeball the data in scatter plots to determine if the data points show a positive association, a negative association, or neither. Lines fitted to data showing positive association have positive slope, and lines fitted to data showing negative association have negative slope.

If most students struggle with this item, plan to reinforce language about positive and negative slopes, specifically "as the value of one variable increases, the value of the other increases" for positive slope, and "as the value of one variable increases, the value of the other decreases" for negative slope.

Statement

A line contains the points $(-3, -2)$ and $(7, 2)$.

1. Without doing any calculations, determine whether the slope of this line is positive or negative. Explain or show your reasoning.
2. Calculate the slope of the line.

Solution

1. The slope is positive. One way to see this is to plot the two points. Another is to note that as the x -coordinate increases, the y -coordinate also increases.
2. $\frac{2}{5}$ (or equivalent)

Aligned Standards

8.F.B.4

Problem 2

The content assessed in this problem is first encountered in Lesson 6: The Slope of a Fitted Line.

Though this problem does not use the word “slope,” students need to find the rate of change for the situation. If most students struggle with this item, plan to have students show their calculations and perhaps revisit this item during Activity 1.

Statement

Diego has some money in his bank account before he starts a summer job. He deposits the money from the summer job in his bank account and doesn't spend any of it. After working 3 hours, he has \$71 in the account. After working 12 hours, he has \$134 in the account.

How much money does Diego earn per hour?

Solution

Diego earns \$7 per hour. He earns \$63 by working 9 hours, so he earns \$7 by working one hour.

Aligned Standards

6.RP.A.3.b

Problem 3

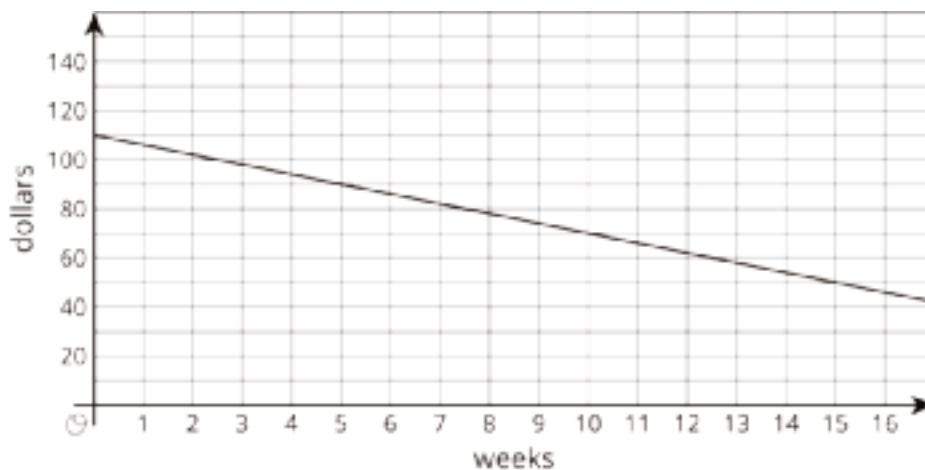
The content assessed in this problem is first encountered in Lesson 8: Analyzing Bivariate Data.

In this unit, students will draw lines that fit data in a scatterplot and estimate the slope, intercepts, and equation for the line they drew.

If most students struggle with this item, beginning in Lesson 4 when students begin connecting the meaning of slope to models, review how slope can be seen on graphs. In Lessons 5, 6, and 7, continue to ask students to show you how the slope and intercepts of any given linear model can be seen in the equation and in the graph. When students draw lines of best fit, give them opportunities to practice estimating slope and intercepts by referring to how slope can be seen on the graph. Help students connect the visual of a slope triangle to the language “for every 1 unit increase in [the independent variable], the [dependent variable] increases/decreases by....” Make sure students attend to the scale of the axes as they estimate and describe slopes.

Statement

Here is a graph showing the balance in someone's savings account since the beginning of the year.



1. Write an equation for the line shown on the graph.
2. What does the slope mean in the situation?
3. What does the vertical intercept mean in the situation?

Solution

1. $d = -4w + 110$ or equivalent
2. The slope of -4 means this person is spending \$4 per week.
3. The vertical intercept of 110 means this person started the year with \$110 in the account.

Aligned Standards

8.EE.B.6, 8.F.B.4

Problem 4

The content assessed in this problem is first encountered in Lesson 9: Looking for Associations.

In this unit, students will be introduced to two-way tables. Many questions that come up in the context of two-way tables are of the form, "What percentage of ____ are ____?"

If most students struggle with this item, plan to spend time during the Launch of Activity 3 helping students verify that the percent of 10 to 12 year olds that have a cell phone is 42%. Students will have more practice finding percents in Lesson 10, but you may choose to revisit this item as part of the Warm-Up for Activity 10.

Statement

In many schools, students have the choice between taking art, music, or some other elective.

1. At Euclid Middle School, there are 200 students in the 8th grade. 40 students are taking art. What percentage of 8th graders at Euclid Middle School are taking art? Explain or show your reasoning.

- 
2. At Newton Middle School, there are 320 students in the 8th grade. 54 are taking music. What percentage of 8th graders at Newton Middle School are taking music? Explain or show your reasoning.

Solution

1. 20%
2. 16.875%

In part a, students might use varying strategies. One strategy is to notice that $100 : 20$ and $200 : 40$ are equivalent ratios. Part b does not yield as easily to mental shortcuts. Students will probably divide 54 by 320 to compute the percentage.

Aligned Standards

6.RP.A.3.c

Problem 5

The content assessed in this problem is first encountered in Lesson 9: Looking for Associations.

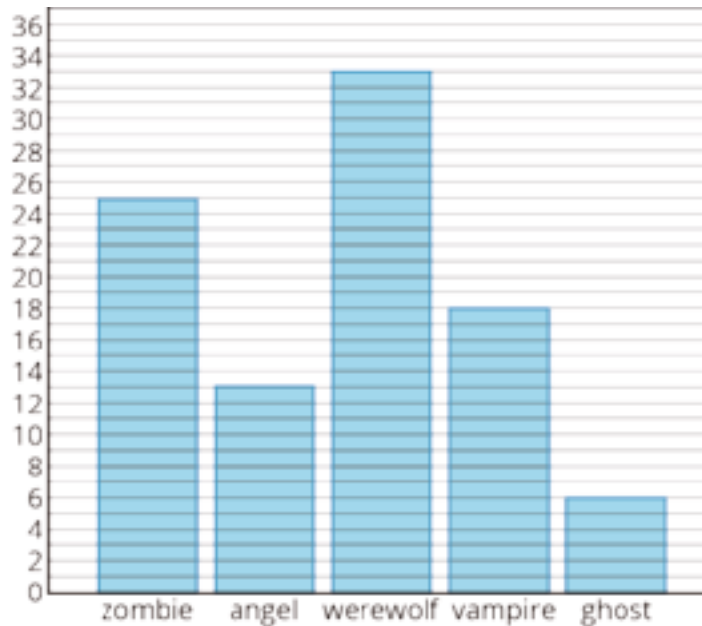
Bar graphs should be familiar to students from grade school, but it has been a long time since students have used this representation. In this unit, students will learn to use segmented bar graphs as a way of representing categorical data.

If most students struggle with this item, plan to ask students in Activity 2 to connect each number in the two-way table to the graph. As students do the card sort in the Activity, you may wish to check in with those who struggled with this item to ask them to map the numbers from the two-way tables to the bar graph it matches.



Statement

Students voted for their favorite entry in a Halloween costume contest. Jada recorded the results in a bar graph:



1. Who got more votes: the angel or the zombie?
2. How many votes did the vampire get?
3. Who won the contest?

Solution

1. The zombie
2. 18 votes
3. The werewolf

Aligned Standards

2.MD.D.10

Problem 6

The content assessed in this problem is first encountered in Lesson 10: Using Data Displays to Find Associations.

Students will have to think about how many bars are needed and how to scale the vertical axis.

If most students struggle with this item, plan to ask students in Activity 1 to work with a partner to create a bar graph from the data. Each partner should use one category as the x-axis label. Partners can trade and check each other's graphs.

Statement

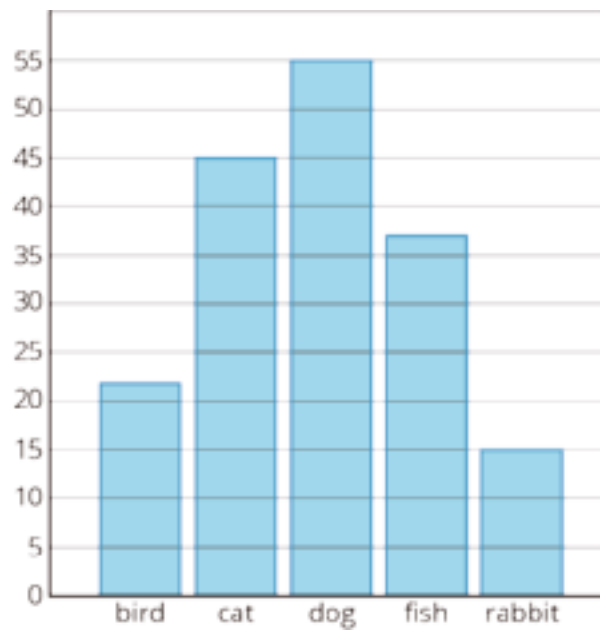
Students at Elena's school were polled about the animal they would most like to have as a pet.

animal	votes
bird	22
cat	45
dog	55
fish	37
rabbit	15

Make a bar graph that displays this information.

Solution

Students should use a scale of 5 or 10 on the vertical axis. Here is one possible bar graph:



Aligned Standards

3.MD.B.3

Assessment : Check Your Readiness (B)

Teacher Instructions

Graphing calculators should not be used. Use of a four-function or scientific calculator is acceptable. Provide access to graph paper.

Student Instructions

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

Problem 1

The content assessed in this problem is first encountered in Lesson 5: Describing Trends in Scatter Plots.

In this unit, students will need to eyeball the data in scatter plots to determine if the data points show a positive association, a negative association, or neither. Lines fitted to data showing positive association have positive slope, and lines fitted to data showing negative association have negative slope.

If most students struggle with this item, plan to reinforce language about positive and negative slopes, specifically "as the value of one variable increases, the value of the other increases" for positive slope, and "as the value of one variable increases, the value of the other decreases" for negative slope.

Statement

A line contains the points $(-1, 4)$ and $(5, -3)$.

1. Without doing any calculations, determine whether the slope of this line is positive or negative. Explain or show your reasoning.
2. Calculate the slope of the line.

Solution

1. The slope is negative. One way to see this is to plot the two points. Another is to note that as the x-coordinate increases, the y-coordinate decreases.
2. $-\frac{7}{6}$ (or equivalent)

Aligned Standards

8.F.B.4

Problem 2

The content assessed in this problem is first encountered in Lesson 6: The Slope of a Fitted Line.

Though this problem does not use the word “slope,” students need to find the rate of change for the situation. If most students struggle with this item, plan to have students show their calculations and perhaps revisit this item during Activity 1.

Statement

Lin has some money in her bank account before she starts a summer job. She deposits the money from the summer job in her bank account and doesn't spend any of it. After working 4 hours, she has \$85 in the account. After working 15 hours, she has \$217 in the account. How much money does Lin earn per hour?

Solution

Lin earns \$12 per hour. She earns \$132 by working 11 hours, so she earns \$12 by working one hour.

Aligned Standards

6.RP.A.3.b

Problem 3

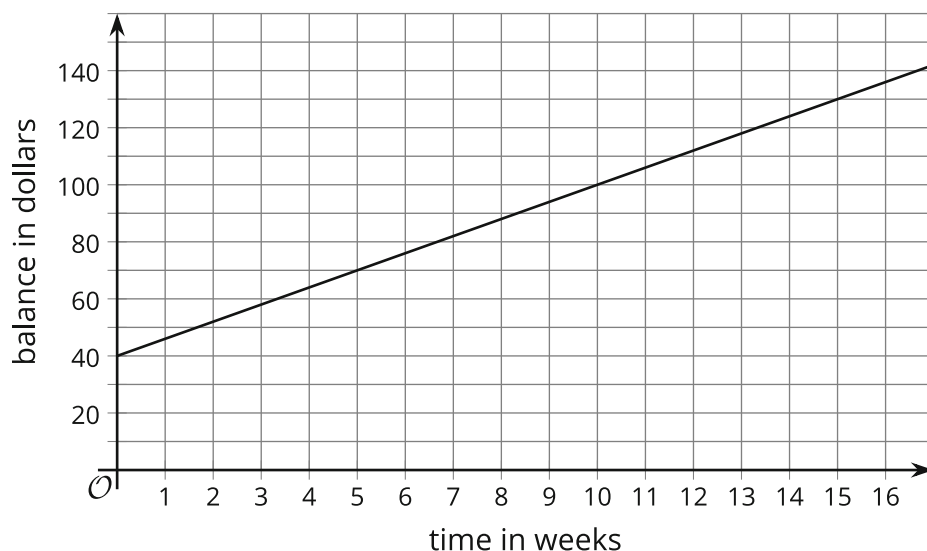
The content assessed in this problem is first encountered in Lesson 8: Analyzing Bivariate Data.

In this unit, students will draw lines that fit data in a scatterplot and estimate the slope, intercepts, and equation for the line they drew.

If most students struggle with this item, beginning in Lesson 4 when students begin connecting the meaning of slope to models, review how slope can be seen on graphs. In Lessons 5, 6, and 7, continue to ask students to show you how the slope and intercepts of any given linear model can be seen in the equation and in the graph. When students draw lines of best fit, give them opportunities to practice estimating slope and intercepts by referring to how slope can be seen on the graph. Help students connect the visual of a slope triangle to the language “for every 1 unit increase in [the independent variable], the [dependent variable] increases/decreases by....” Make sure students attend to the scale of the axes as they estimate and describe slopes.

Statement

Here is a graph showing the balance in someone's savings account since the beginning of the year.



1. Write an equation for the line shown on the graph.
2. What does the slope mean in the situation?
3. What does the vertical intercept mean in the situation?

Solution

1. $b = 6t + 40$ or equivalent
2. The slope of 6 means the balance in the account is increasing by \$6 each week.
3. The vertical intercept of 40 means this person started the year with \$40 in the account.

Aligned Standards

8.EE.B.6, 8.F.B.4

Problem 4

The content assessed in this problem is first encountered in Lesson 9: Looking for Associations.

In this unit, students will frequently calculate percentages, and should be capable of rounding to the nearest percent or tenth of a percent.

If most students struggle with this item, plan to spend time during the Launch of Activity 3 helping students verify that the percent of 10 to 12 year olds that have a cell phone is 42%. Students will have more practice finding percents in Lesson 10, but you may choose to revisit this item as part of the Warm-Up for Activity 10.

Statement

Determine each answer to the nearest tenth of a percent.

1. 8 out of 32 football teams are on the East Coast. What percentage of football teams are on the East Coast?
2. It rained 4 out of 7 days this week. What percentage of days did it rain?

Solution

1. 25.0%
2. 57.1%

Aligned Standards

6.RP.A.3.c

Problem 5

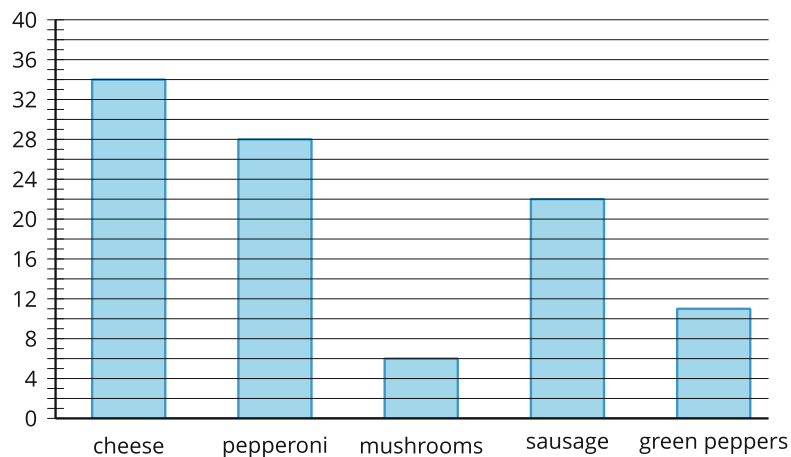
The content assessed in this problem is first encountered in Lesson 9: Looking for Associations.

Bar graphs should be familiar to students from grade school, but it has been a long time since students have used this representation. In this unit, students will learn to use segmented bar graphs as a way of representing categorical data.

If most students struggle with this item, plan to ask students in Activity 2 to connect each number in the two-way table to the graph. As students do the card sort in the Activity, you may wish to check in with those who struggled with this item to ask them to map the numbers from the two-way tables to the bar graph it matches.

Statement

Students voted for their favorite pizza toppings. Kiran recorded the results in a bar graph.



1. Which topping won the contest?
2. How many votes did pepperoni get?
3. Which topping got more votes: mushrooms or green peppers?

Solution

1. cheese
2. 28
3. green peppers

Aligned Standards

2.MD.D.10

Problem 6

The content assessed in this problem is first encountered in Lesson 10: Using Data Displays to Find Associations.

Students will have to think about how many bars are needed and how to scale the vertical axis.

If most students struggle with this item, plan to ask students in Activity 1 to work with a partner to create a bar graph from the data. Each partner should use one category as the x-axis label. Partners can trade and check each other's graphs.

Statement

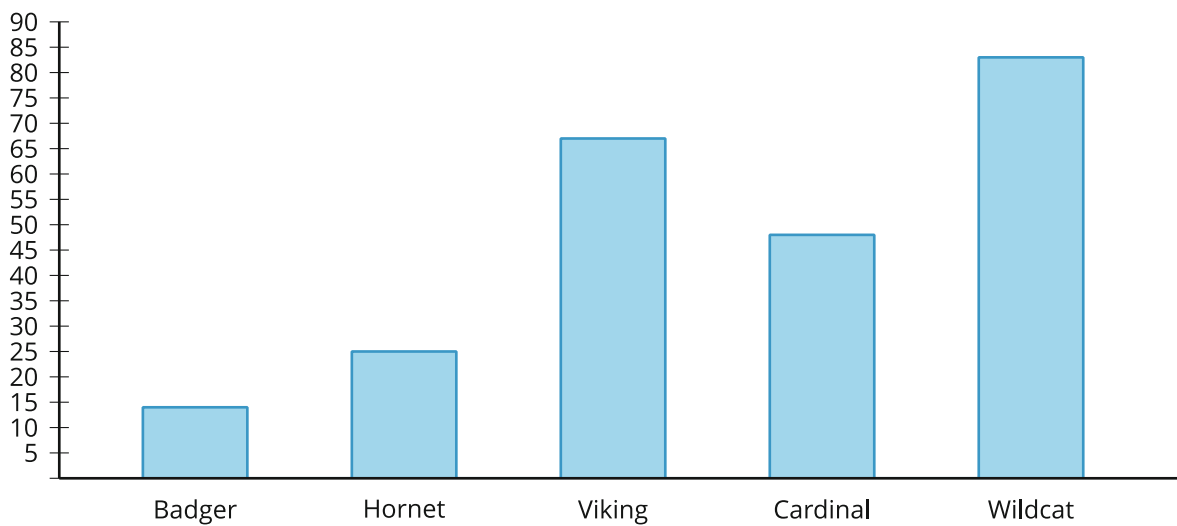
Students at Andre's school were polled about the mascot they would most like to have for their school.

mascot	number of votes
badger	14
hornet	25
viking	67
cardinal	49
wildcat	83

Make a bar graph that displays this information.

Solution

Students should use a scale of 5 or 10 on the vertical axis. Here is one possible bar graph:



Aligned Standards

3.MD.B.3

Assessment : End-of-Unit Assessment (A)

Teacher Instructions

Graphing calculators should not be used. Use of a four-function or scientific calculator is acceptable. Provide access to graph paper.

Student Instructions

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

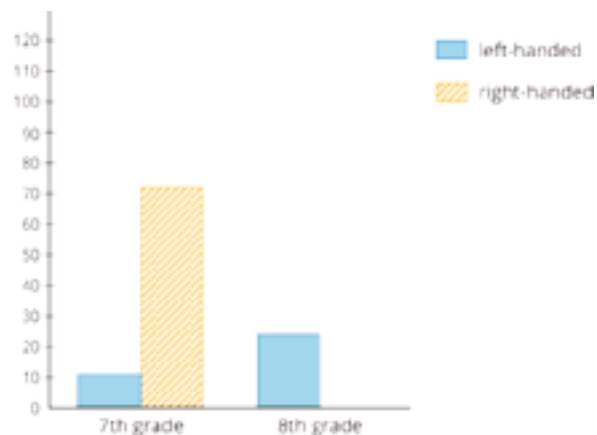
Problem 1

Students selecting A have reversed the proportion between the columns ($72 \cdot \frac{11}{24}$). Students selecting B are preserving the absolute difference in each column. Students selecting C have added the numbers in the problem: $(11 + 24 + 72)$.

Statement

Noah gathered data at his school among 7th and 8th graders to see if there was an association between grade level and handedness. The table and graph show his data, but the number of right-handed 8th graders is missing.

	left-handed	right-handed
7th grade	11	72
8th grade	24	



Noah found there was *no evidence* of an association between grade level and handedness. Which of these could be the number of right-handed 8th graders?

- A. 33
- B. 85
- C. 107
- D. 157

Solution

D

Aligned Standards

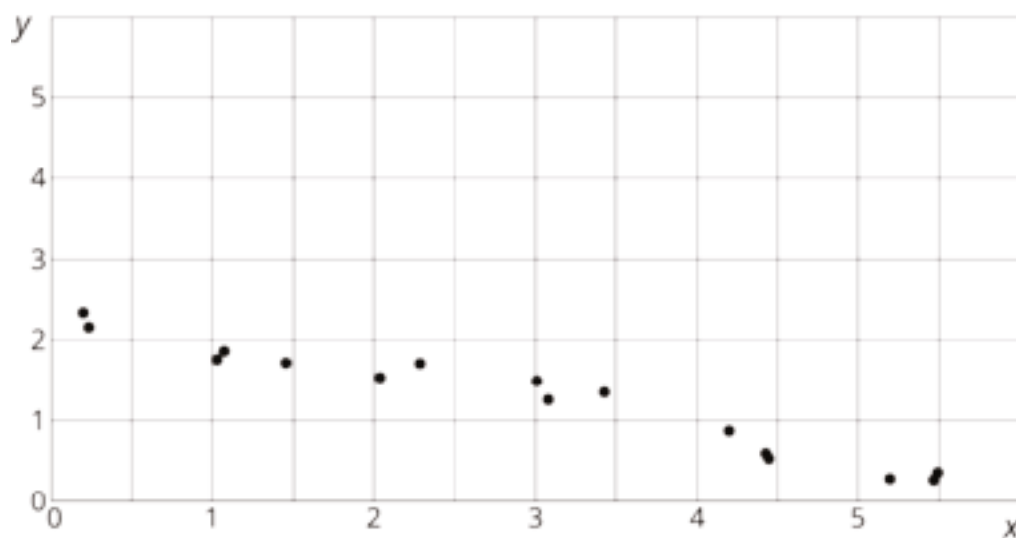
8.SP.A.4

Problem 2

Students selecting B have made an error in picking the intercept, using $(6, 0)$ instead of $(0, 2)$. Students selecting C have the intercept right but have incorrectly chosen the sign of the slope. Students selecting D have made both of these errors.

Statement

Here is a scatter plot:



The graph of what linear equation is a good fit for this data?

A. $y = \frac{-1}{3}x + 2$

B. $y = \frac{-1}{3}x + 6$

C. $y = \frac{1}{3}x + 2$

D. $y = \frac{1}{3}x + 6$

Solution

A

Aligned Standards

8.SP.A.2

Problem 3

A student selecting A may not be thinking about the way outside temperature is interpreted; as temperature increases, the cost to heat a home generally decreases. Students failing to select C or B have missed relatively clear-cut associations. A student selecting choice D may be looking for any association, without regard to whether it is positive or negative. A student failing to select E may be thinking about a required linear relationship, instead of a generally positive association.

Statement

Select **all** the relationships that demonstrate a positive association between variables.

- A. Outside temperature and cost to heat a home
- B. Length of time you have walked and distance you have traveled
- C. Pounds of cherries you buy and amount of money you spend on cherries
- D. Speed of a train and the amount of time it takes for the train to get to its destination
- E. Number of people in a grocery check-out line and how long you have to wait to check out

Solution

["B", "C", "E"]

Aligned Standards

8.SP.A.1

Problem 4

This problem asks students to demonstrate their understanding of the terms “positive association,” “negative association,” and “outlier” by drawing scatter plots. The points need not be arranged close to a line to show an association: the important thing is that there is a general upward (or downward) trend.

Statement

1. Draw a scatter plot that shows a positive, linear association and has one clear outlier. Circle the outlier.
2. Draw a scatter plot that shows a negative association that is not linear.

Solution

Answers vary.

Minimal Tier 1 response:

- Work is complete and correct.

- Sample:

1. Plot shows 3 points nearly in the same line with a positive slope, and 1 circled point not near the line.
2. Plot shows 3 points that are not nearly in the same line, with a generally negative trend.

Tier 2 response:

- Work shows general conceptual understanding and mastery, with some errors.
- Sample errors: one plot shows a significant error, but the other is correct; outlier not circled; plots have reversed negative and positive associations, but otherwise all work is correct.

Tier 3 response:

- Significant errors in work demonstrate lack of conceptual understanding or mastery.
- Sample errors: both plots show errors beyond the ones listed in Tier 2.

Aligned Standards

8.SP.A.1

Problem 5

This problem requires students to understand the structure of a two-way table. Successful students will subtract across the rows to find the missing entries in the first two rows, then add down the columns for the bottom row.

Statement

Jada surveyed all 7th and 8th graders at her school about whether they have pets. Complete the missing entries in this two-way table.

	has pet	has no pet	total
7th grade	102		150
8th grade		68	175
total			

Solution

	has pet	has no pet	total
7th grade	102	48	150
8th grade	107	68	175
total	209	116	325

Aligned Standards

8.SP.A.4

Problem 6

This problem assesses understanding of relative frequency, both by having students calculate relative frequency from a two-way table and by constructing segmented bar graphs to represent the data.

Statement

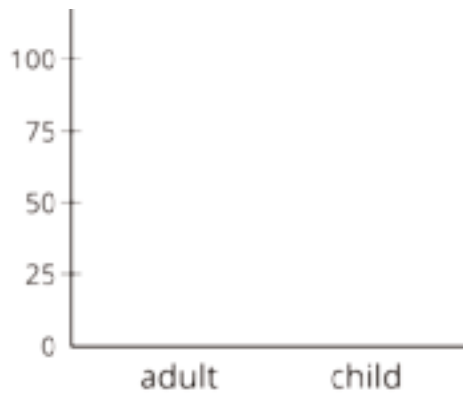
At a school social, children attend with family members. Everyone had a choice between a sweet snack and a salty snack. Here is a two-way table showing the number of adults and children who made each choice of snack.

	sweet snack	salty snack	total
adult	57	88	145
child	77	31	108
total	134	119	253

1. Complete the table with relative frequency by row. Round to the nearest percent.

	sweet snack	salty snack	total
adult			100%
child			100%

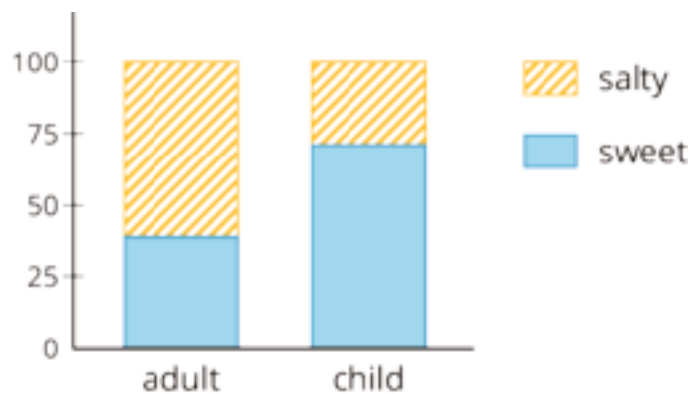
2. Make a segmented bar graph to represent the data in your table. Use one bar for each row of the table.



Solution

1.

	sweet snack	salty snack	total
adult	39%	61%	100%
child	71%	29%	100%



Minimal Tier 1 response:

- Work is complete and correct.
- Sample:
 1. The table is completely correct.
 2. The segmented bar graphs are accurate.

Tier 2 response:

- Work shows general conceptual understanding and mastery, with some errors.
- Sample errors: percentages in table are nearly correct but wrong due to a calculation or rounding error; segmented bar graphs are inaccurate.

- Acceptable errors: segmented bar graph is correct based on incorrect percentages.

Tier 3 response:

- Significant errors in work demonstrate lack of conceptual understanding or mastery.
- Sample errors: a regular bar graph is drawn instead of a segmented bar graph; percentages in table are wildly inaccurate, notably if outside 0-100% range or if using an incorrect total as the basis for the percentages.

Aligned Standards

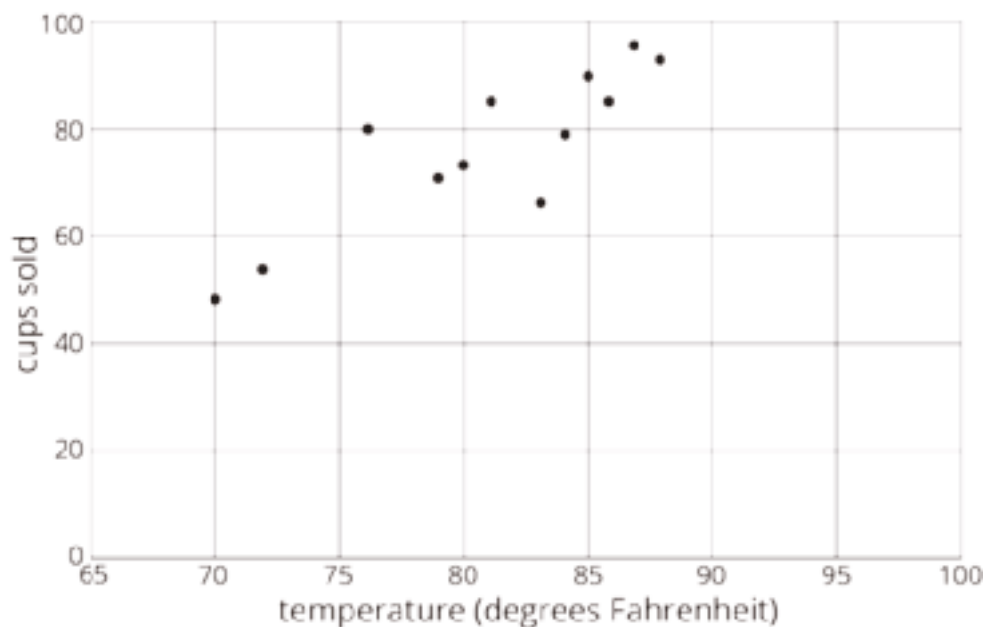
8.SP.A.4

Problem 7

In part c, students could also solve this problem by choosing a temperature for Saturday, such as 70 degrees, then using the equation $c = 2t - 89$ to calculate the expected number of cups sold for 70 degrees and for 75 degrees, then comparing the answers.

Statement

Lin opened a lemonade stand during the summer. She noticed that she sold more lemonade on warmer days. For each day she sold lemonade, she plotted the point (t, c) , where t represents high temperature and c represents cups of lemonade sold.

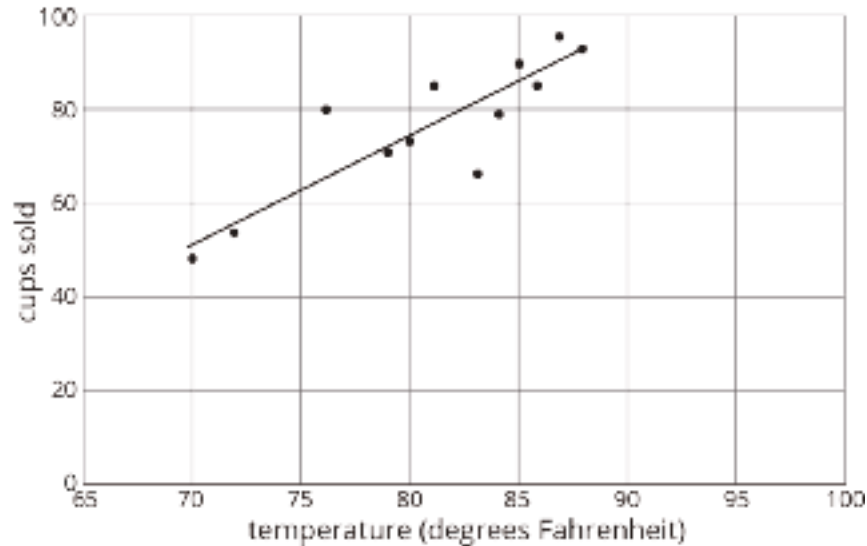


1. On the same axes, draw a line that you think is a good fit for the data.
2. A computer program found that the line $c = 2t - 89$ is a good fit for the data. Use this equation to predict how many cups of lemonade Lin might sell on a day when the high temperature is 74 degrees.

3. The high temperature this Sunday is expected to be 5 degrees warmer than the high temperature this Saturday. Using the line $c = 2t - 89$, how many more cups of lemonade should Lin expect to sell on Sunday than Saturday? Explain or show your reasoning.

Solution

1. Answers vary. Sample response:



- b. 59 cups. $2 \cdot 74 - 89 = 59$.

- c. 10 more cups. The slope of the line is 2. This means that for each one-degree increase in temperature, Lin can expect to sell about two more cups of lemonade. If the temperature increases by 5 degrees, she can expect to sell about 10 more cups of lemonade.

Minimal Tier 1 response:

- Work is complete and correct, with complete explanation or justification.
- Sample:
 1. The line reasonably represents a fitting line for the data.
 2. 59
 3. 10, because the slope is 2 and $2 \cdot 5 = 10$.

Tier 2 response:

- Work shows good conceptual understanding and mastery, with either minor errors or correct work with insufficient explanation or justification.

- Sample errors: fitting line is not close but still generally follows direction of data; minor visible calculation error gives incorrect number of cups for 74 degrees; explanation missing for why 10 more cups are expected on Sunday.

Tier 3 response:

- Work shows a developing but incomplete conceptual understanding, with significant errors.
- Sample errors: badly incorrect or missing fitting line; incorrect answer given for 74-degree problem with no work shown or explanation; incorrect application of rule $c = 2t - 89$.

Tier 4 response:

- Work includes major errors or omissions that demonstrate a lack of conceptual understanding and mastery.
- Sample errors: two or more error types from Tier 3 response.

Aligned Standards

8.SP.A.2, 8.SP.A.3

Assessment : End-of-Unit Assessment (B)

Teacher Instructions

Graphing calculators should not be used. Use of a four-function or scientific calculator is acceptable. Provide access to graph paper.

Student Instructions

You may use graph paper and a four-function or scientific calculator, but not a graphing calculator.

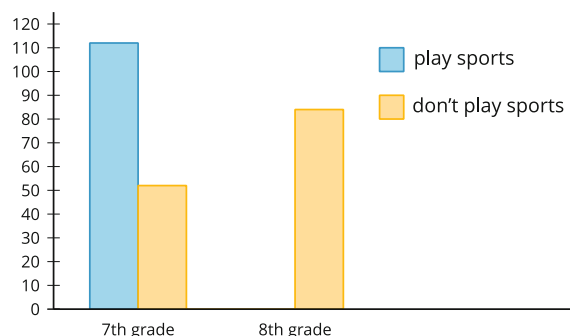
Problem 1

Students selecting A have added the numbers in the problem: $(112 + 52 + 84 = 248)$. Students selecting C subtracted $84 - 52 = 32$. Then added $32 + 112 = 144$. Students selecting D have reversed the proportion between the columns $(112 \cdot \frac{52}{84} = 69)$.

Statement

Jada gathered data at her school among 7th and 8th graders to see if there was an association between grade level and playing sports. The table shows her data, but the number of 8th graders who play sports is missing.

	play sports	don't play sports
7th graders	112	52
8th graders		84



Jada found there was *no evidence* of an association between grade level and playing sports. Which of these could be the number of 8th graders who play sports?

- A. 248
- B. 181
- C. 144
- D. 69

Solution

B

Aligned Standards

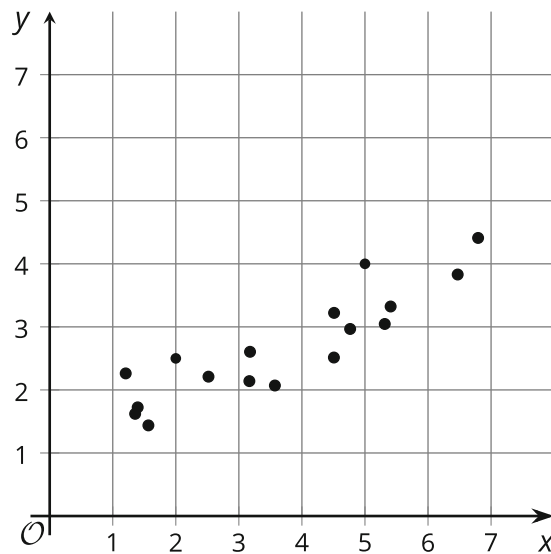
8.SP.A.4

Problem 2

Students selecting A have made an error in picking the intercept, using (0,0) instead of (0,1) .
Students selecting B have the intercept right but have incorrectly chosen the sign of the slope.
Students selecting C have made both of these errors.

Statement

Here is a scatter plot:



The graph of what linear equation is a good fit for this data?

- A. $y = \frac{1}{2}x$
- B. $y = -\frac{1}{2}x + 1$
- C. $y = -\frac{1}{2}x$
- D. $y = \frac{1}{2}x + 1$

Solution

D

Aligned Standards

8.SP.A.2

Problem 3

Students failing to select A, C, and E may be thinking about a required linear relationship, instead of a generally negative association. Students failing to select A, C or E have missed relatively clear-cut

associations. Students selecting choice B and D may be looking for any association, without regard to whether it is positive or negative.

Statement

Select **all** the relationships that demonstrate a negative association between variables.

- A. Number of absences from school and final grades
- B. Outside temperature and ice cream sales
- C. Price of houses and house sales
- D. Number of rainy days and car accidents
- E. Number of hours playing video games and grades

Solution

["A", "C", "E"]

Aligned Standards

8.SP.A.1

Problem 4

This problem asks students to demonstrate their understanding of the terms “positive association,” “negative association,” and “outlier” by drawing scatter plots. The points need not be arranged close to a line to show an association: the important thing is that there is a general upward (or downward) trend.

Statement

1. Draw a scatter plot that shows a negative, linear association and has one clear outlier. Circle the outlier.
2. Draw a scatter plot that shows a positive association that is not linear.

Solution

Answers vary.

Minimal Tier 1 response:

- Work is complete and correct.
- Sample:
 1. Plot shows 3 points nearly in the same line with a negative slope, and 1 circled point not near the line.
 2. Plot shows 3 points that are not nearly in the same line, with a generally positive trend.

Tier 2 response:

- Work shows general conceptual understanding and mastery, with some errors.
- Sample errors: one plot shows a significant error, but the other is correct; outlier not circled; plots have reversed negative and positive associations, but otherwise all work is correct.

Tier 3 response:

- Significant errors in work demonstrate lack of conceptual understanding or mastery.
- Sample errors: both plots show errors beyond the ones listed in Tier 2.

Aligned Standards

8.SP.A.1

Problem 5

This problem requires students to understand the structure of a two-way table. Successful students will subtract across the rows to find the missing entries in the first two rows, then add down the columns for the bottom row.

Statement

Andre surveyed all 7th and 8th grade students at his school about whether they bring their lunch or eat the school lunch. Complete the missing entries in this two-way table.

	bring home lunch	eat school lunch	total
7th grade students		219	293
8th grade students	168		261
total			

Solution

Here is the completed table:

	bring home lunch	eat school lunch	total
7th grade students	74	219	293
8th grade students	168	93	261
total	242	312	554

Aligned Standards

8.SP.A.4

Problem 6

This problem assesses understanding of relative frequency, both by having students calculate relative frequency from a two-way table and by constructing segmented bar graphs to represent the data.

Statement

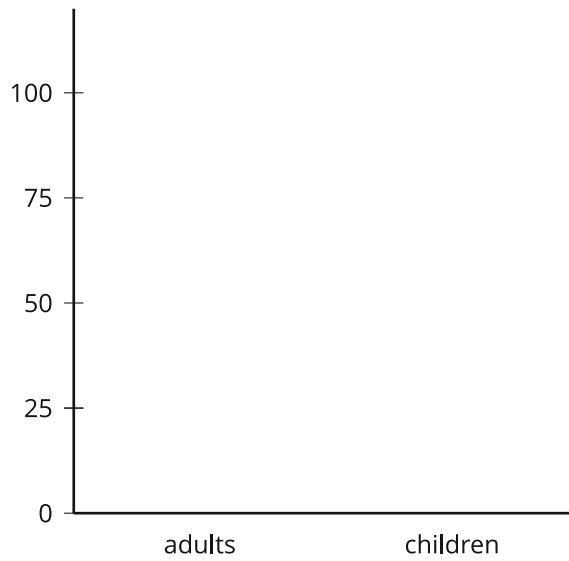
At the school carnival, adults attended with children. Everyone had a choice between a piece of pizza and a hot dog. Here is a two-way table showing the number of adults and children who made each choice.

	pizza	hot dog	total
adult	48	105	153
child	92	73	165
total	140	178	318

1. Complete the table with relative frequency by row. Round to the nearest percent.

	pizza	hot dog	total
adult			100%
child			100%

2. Make a segmented bar graph to represent the data in your table. Use one bar for each row of the table.

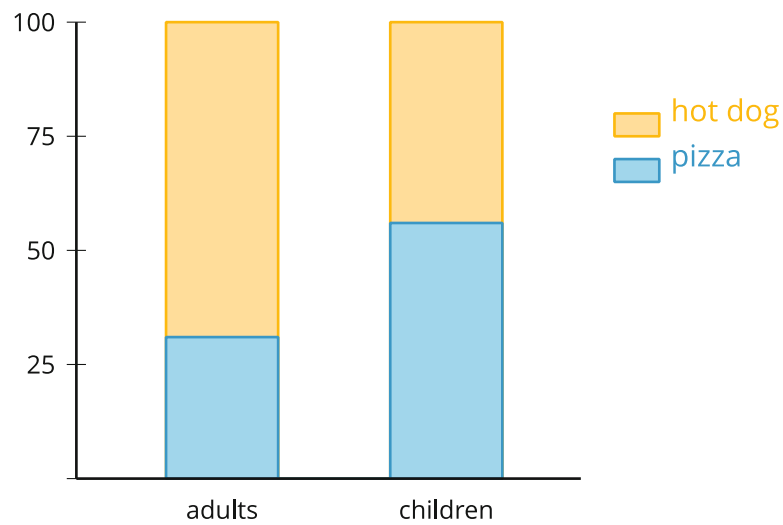


Solution

1.

	pizza	hot dog	total
adult	31%	69%	100%
child	56%	44%	100%

2.



Minimal Tier 1 response:

- Work is complete and correct.

- Sample:
- The table is completely correct.
- The segmented bar graphs are accurate.

Tier 2 response:

- Work shows general conceptual understanding and mastery, with some errors.
- Sample errors: percentages in table are nearly correct but wrong due to a calculation or rounding error; segmented bar graphs are inaccurate.
- Acceptable errors: segmented bar graph is correct based on incorrect percentages.

Tier 3 response:

- Significant errors in work demonstrate lack of conceptual understanding or mastery.
- Sample errors: a regular bar graph is drawn instead of a segmented bar graph; percentages in table are wildly inaccurate, notably if outside 0-100% range or if using an incorrect total as the basis for the percentages.

Aligned Standards

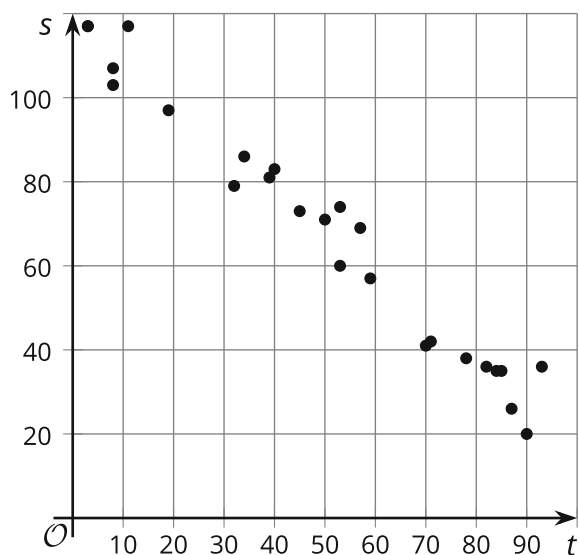
8.SP.A.4

Problem 7

In part c, students could also solve this problem by choosing a temperature for Saturday, such as 40 degrees, then using the equation $s = -t + 120$ to calculate the expected number of bowls sold for 40 degrees and for 30 degrees, then comparing the answers.

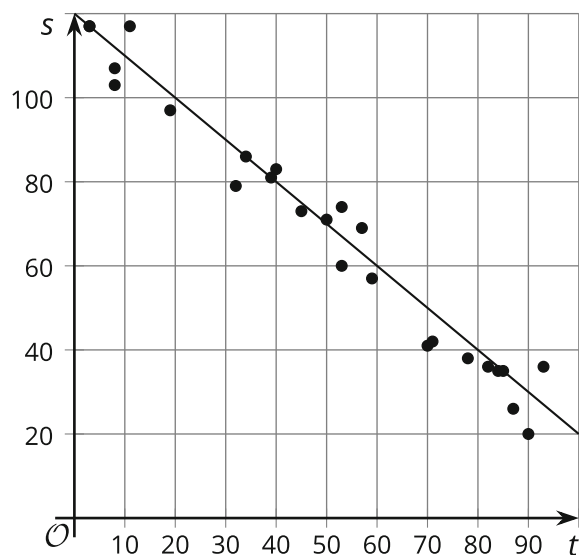
Statement

A deli owner noticed that as the outside temperature increased, they sold less soup. For each day soup was sold, they plotted the point (t, s) , where t represents high temperature and s represents bowls of soup sold.



1. On the same axes, draw a line that you think is a good fit for the data.
2. The deli owner found that the line $s = -t + 120$ is a good fit for the data. Use this equation to predict how many bowls of soup they might sell on a day when the high temperature is 32 degrees.
3. The high temperature this Saturday is expected to be 10 degrees colder than the high temperature this Friday. Using the line $s = -t + 120$, how many more bowls of soup should the deli expect to sell on Saturday than Friday? Explain or show your reasoning.

Solution



1. Answers vary. Sample response:
2. 88 bowls. $-32 + 120 = 88$

3. 10 more bowls. The slope of the line is -1. This means that for each one-degree decrease in temperature, the deli can expect to sell about one more bowl of soup. If the temperature decreases by 10 degrees, they can expect to sell about 10 more bowls of soup.

Minimal Tier 1 response:

- Work is complete and correct, with complete explanation or justification.
- Sample:
 1. The line reasonably represents a fitting line for the data.
 2. 88
 3. 10, because the slope is -1 and $-1 \cdot -10 = 10$

Tier 2 response:

- Work shows good conceptual understanding and mastery, with either minor errors or correct work with insufficient explanation or justification.
- Sample errors: fitting line is not close but still generally follows direction of data; minor visible calculation error gives incorrect number of bowls for 32 degrees; explanation missing for why 10 more bowls are expected on Saturday.

Tier 3 response:

- Work shows a developing but incomplete conceptual understanding, with significant errors.
- Sample errors: badly incorrect or missing fitting line; incorrect answer given for 32-degree problem with no work shown or explanation; incorrect application of rule $s = -t + 120$.

Tier 4 response:

- Work includes major errors or omissions that demonstrate a lack of conceptual understanding and mastery.
- Sample errors: two or more error types from Tier 3 response.

Aligned Standards

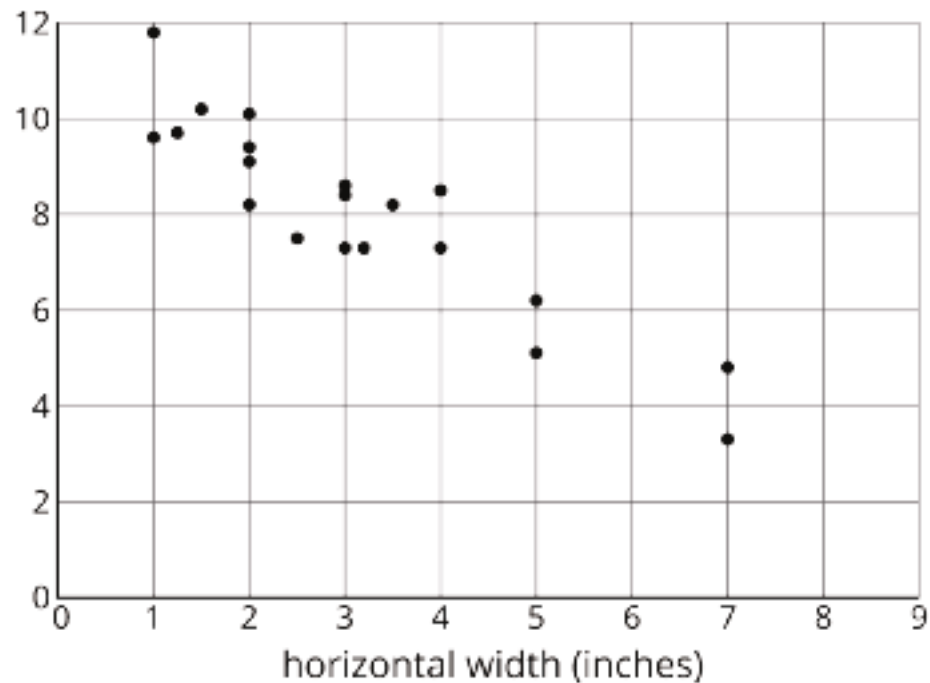
8.SP.A.2, 8.SP.A.3

Lesson Cool Downs

Lesson 1: Organizing Data

Cool Down: Squashed Spheres

Twenty rubber spheres were compressed with varying amounts of force. The widths and heights of the resulting shapes were measured. Here is a scatter plot that shows the measurements for each sphere.



1. Label the vertical axis of the scatter plot.
2. If a compressed rubber sphere has a 6-inch width, is its height closer to 5 inches or to 11 inches? Explain your reasoning.

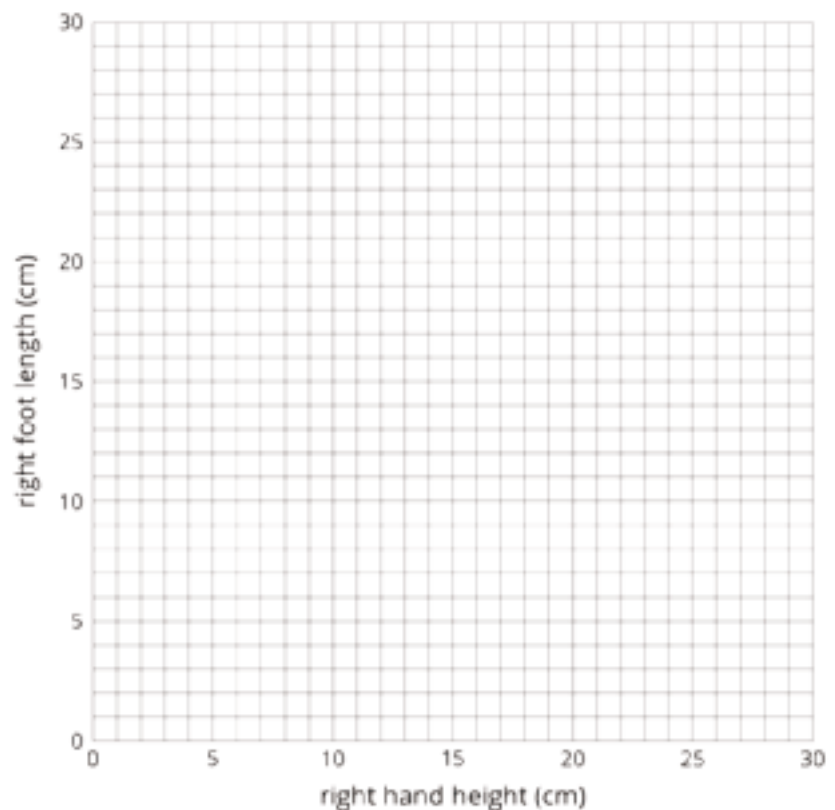
Lesson 2: Plotting Data

Cool Down: Right Side Measurements

The table shows measurements of right hand length and right foot length for five people.

	right hand length (cm)	right foot length (cm)
person A	19	27
person B	21	30
person C	17	23
person D	18	24
person E	19	26

1. Draw a scatter plot for the data.
2. Circle the point in the scatter plot that represents Person D's measurements.

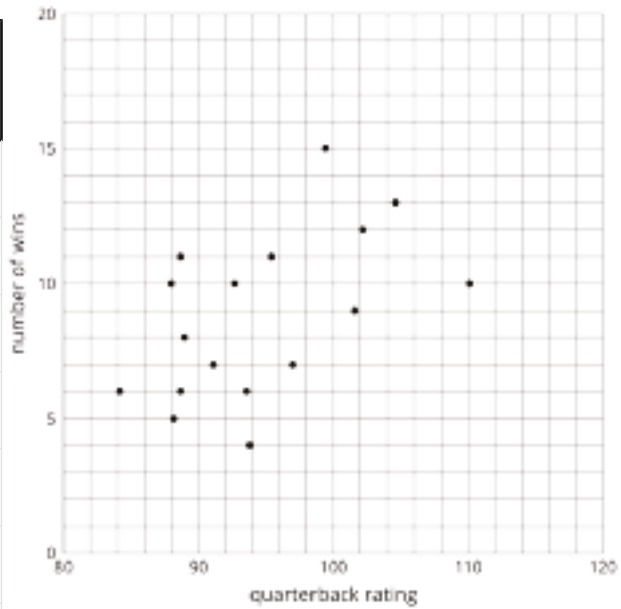


Lesson 3: What a Point in a Scatter Plot Means

Cool Down: Quarterbacks

Here are a table and scatter plot that show ratings and wins for quarterbacks who started 16 games this season.

player	quarterback rating	number of wins
A	93.8	4
B	102.2	12
C	93.6	6
D	89	8
E	88.2	5
F	97	7
G	88.7	6
H	91.1	7
I	92.7	10
J	88	10
K	101.6	9
L	104.6	13
M	84.2	6
N	99.4	15
O	110.1	10
P	95.4	11
Q	88.7	11

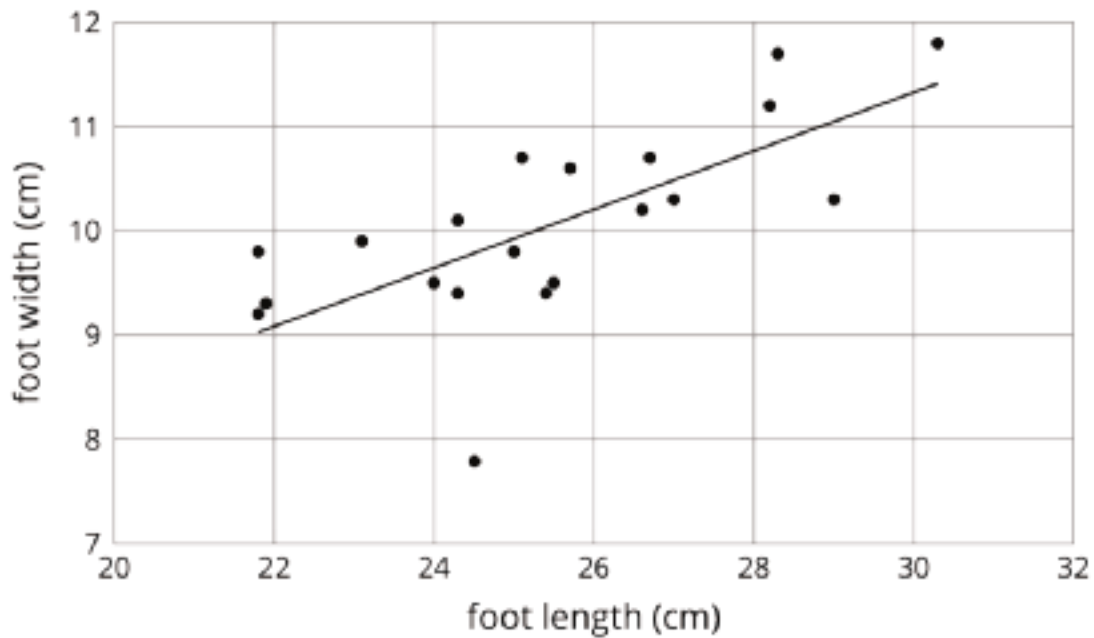


1. Circle the point in the scatter plot that represents Player K's data.
2. Which quarterback's data are represented by the point farthest to the left?
3. Player R is not included in the table because he did not start 16 games this year. He did have a quarterback rating of 99.4 and his team won 8 games. On the scatter plot, plot a point that represents Player R's data.

Lesson 4: Fitting a Line to Data

Cool Down: A 1 Foot Foot

Here is a scatter plot that shows lengths and widths of 20 left feet, together with the graph of a model of the relationship between foot length and width.

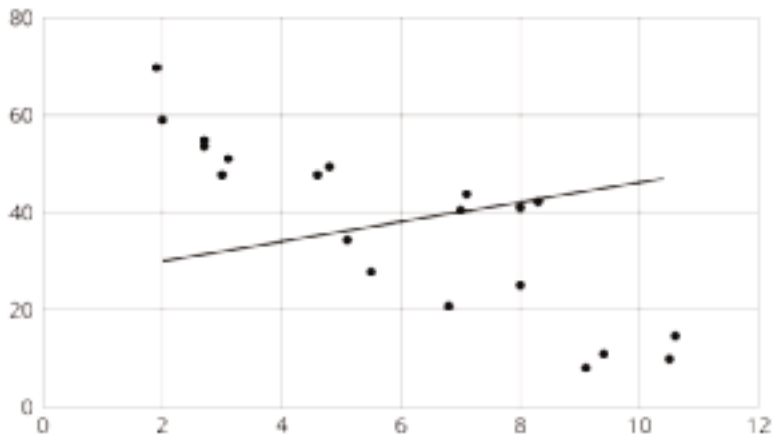


1. Draw a box around the point that represents the foot with length closest to 29 cm.
2. What is the approximate width of this foot?
3. What width does the model predict for a foot with length 29 cm?

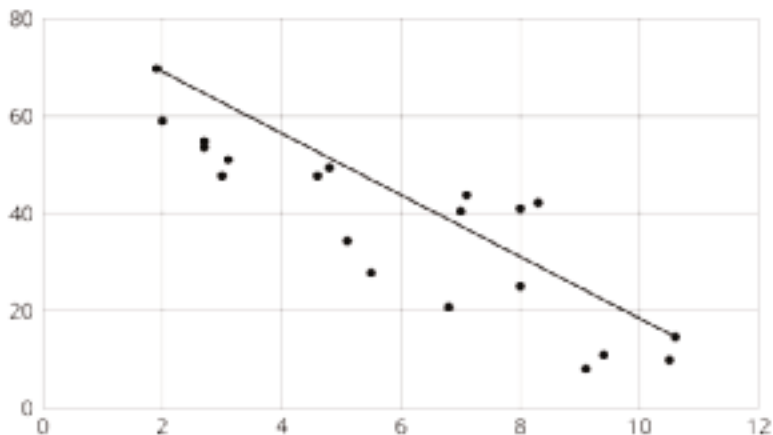
Lesson 5: Describing Trends in Scatter Plots

Cool Down: This is One Way to Do It

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.



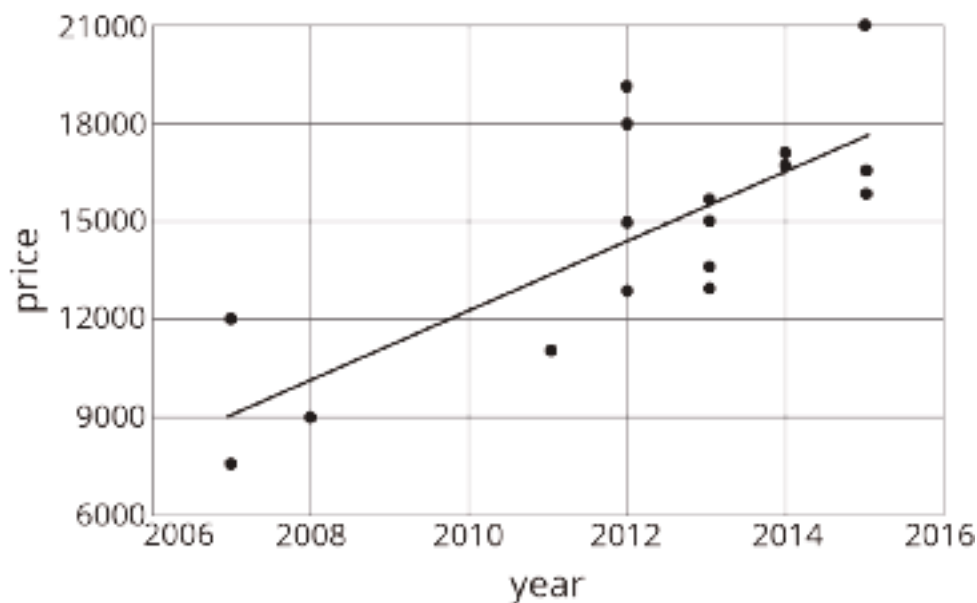
2. Noah said, "I think this line is a good fit because it passes through the leftmost point and the rightmost point." Do you agree? Explain your reasoning.



Lesson 6: The Slope of a Fitted Line

Cool Down: Trends in the Price of Used Cars

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 slope positive or negative?
2. Which of these values is closest to the slope 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 slope in this context.

Lesson 7: Observing More Patterns in Scatter Plots

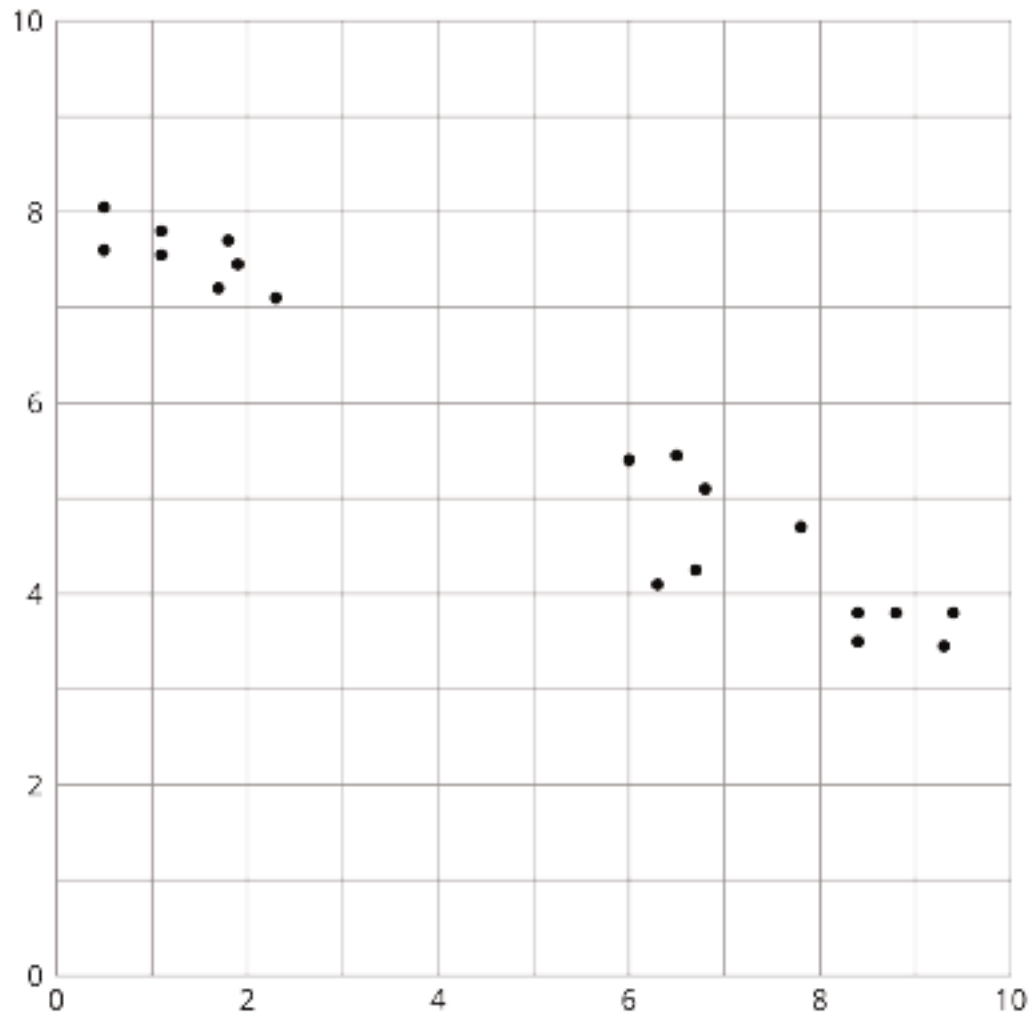
Cool Down: Make Your Own Scatter Plot

1. Draw a scatter plot that shows a positive linear association and clustering.
2. Draw a scatter plot that shows a negative non-linear association and no clustering.

Lesson 8: Analyzing Bivariate Data

Cool Down: Drawing a Line

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



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?

Lesson 9: Looking for Associations

Cool Down: Guitar and Golf

1. In a class of 25 students, some students play a sport, some play a musical instrument, some do both, some do neither. Complete the two-way table to show data that might come from this class.

	plays an instrument	does not play an instrument	total
plays a sport			
does not play a sport	5		
total			25

2. Using the entries from the previous table, complete this table so that it shows relative frequencies. Round entries to the nearest percentage point.

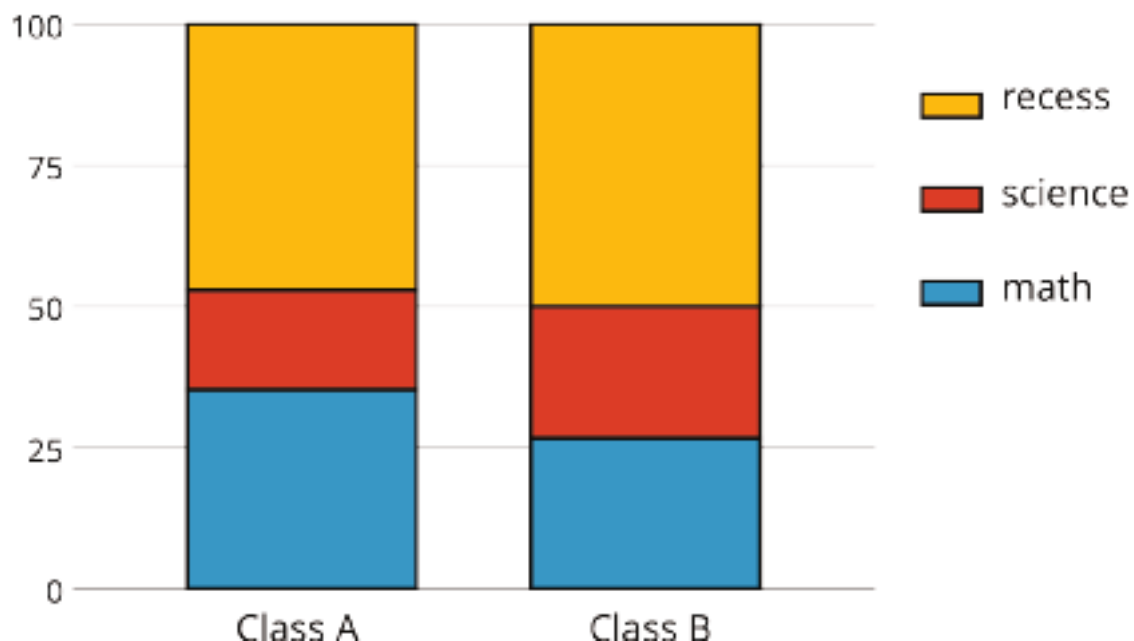
	plays an instrument	does not play an instrument	total
plays a sport			
does not play a sport			

Lesson 10: Using Data Displays to Find Associations

Cool Down: Class Preferences

Here are a two-way table and segmented bar graph for data from students in two classes. Do they show evidence of differences between the two classes?

	prefers math	prefers science	prefers recess
class A	6	3	8
class B	8	7	15



Instructional Masters

Instructional Masters for Associations in Data

address	title	students per copy	written on?	requires cutting?	card stock recommended?	color paper recommended?
Activity Grade8.6.7.2	Scatter Plot City	1	no	yes	no	yes
Activity Grade8.6.1.3	Tables and Their Scatter Plots	1	no	no	no	no
Activity Grade8.6.9.2	Card Sort: Matching Representations	1	no	yes	no	no

20 New Cars

mass (kilograms)	fuel efficiency (miles per gallon)
1,549	25
1,610	20
1,737	21
1,777	20
1,155	28
1,486	23
1,962	16
2,384	16
1,957	19
2,212	16
1,115	29
2,068	18
1,663	19
2,216	18
1,432	25
1,987	18
1,259	25
1,580	26
1,234	30
1,656	23

30 Days in August

daily high (degrees Fahrenheit)	energy consumed (kilowatt hours)
94	37
86	29
92	29
96	38
97	37
92	34
98	36
100	33
100	38
99	39
99	39
87	29
94	43
96	34
95	35
92	36
93	25
90	26
91	28
93	28
90	30
86	32
83	33
90	27
92	29
96	35
84	20
76	17
88	32
77	34

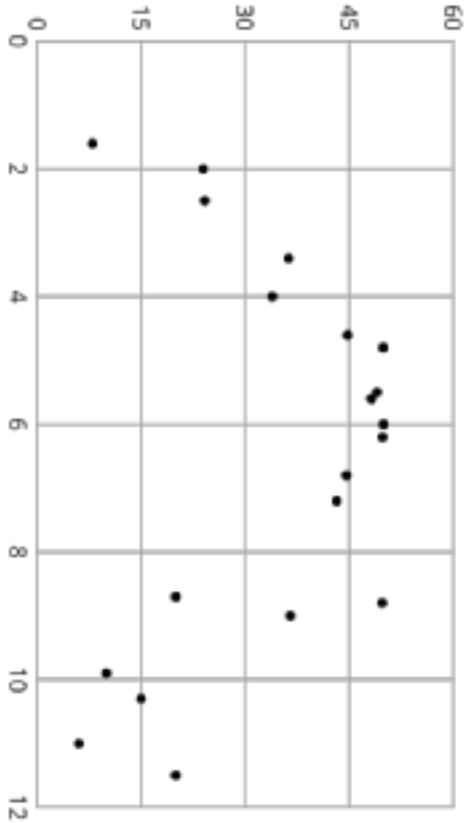
20 Used Cars

price (dollars)	mileage (miles)
7,599	121,000
11,998	70,000
8,998	126,000
8,998	104,000
10,998	69,000
12,998	55,000
12,998	52,000
14,998	23,000
18,998	47,000
17,998	46,000
14,998	23,000
12,998	29,000
13,599	25,000
15,599	2,000
13,599	28,000
16,998	17,000
16,599	12,000
17,000	15,000
15,998	16,000
16,599	9,000
20,998	13,000

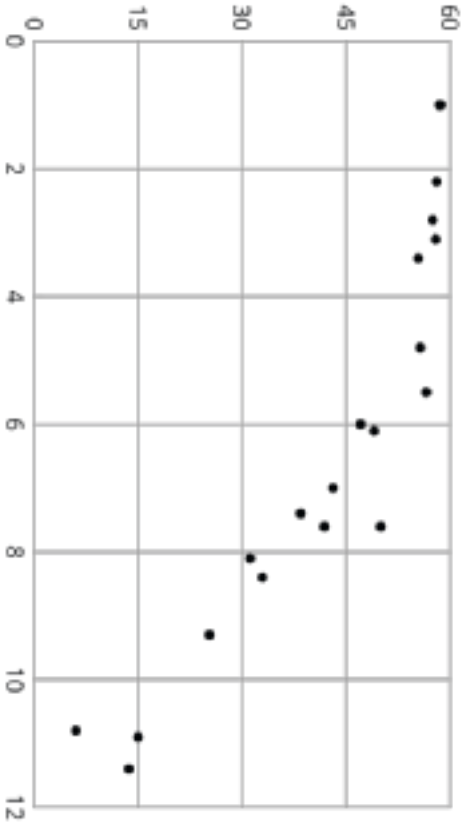
20 Diamonds

weight (carats)	price (dollars)
1	3,772
1	4,221
1	4,032
1	5,385
1.05	3,942
1.05	5,480
1.06	4,511
1.2	5,544
1.3	6,131
1.32	5,872
1.41	7,122
1.5	7,474
1.5	5,904
1.59	8,706
1.61	8,252
1.73	9,530
1.77	9,374
1.85	8,169
1.9	9,541
2.04	9,125

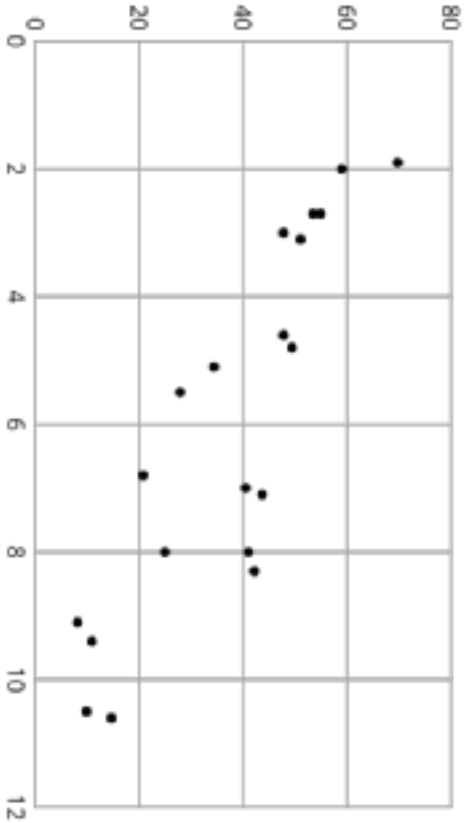
Scatter Plot City
A



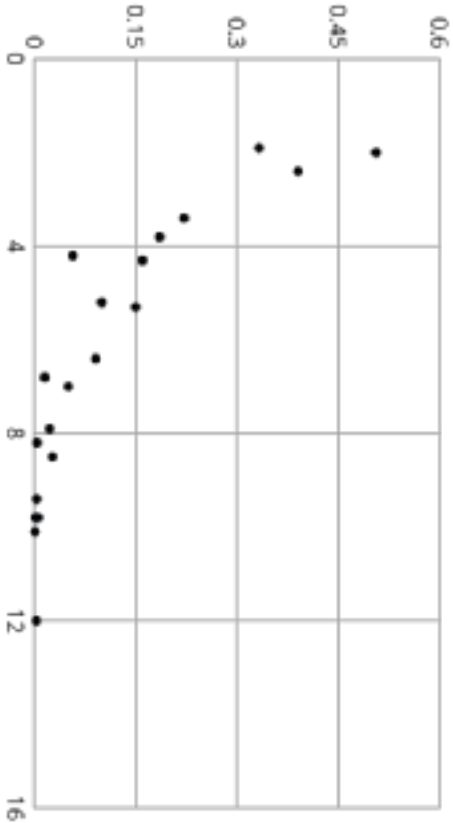
Scatter Plot City
B



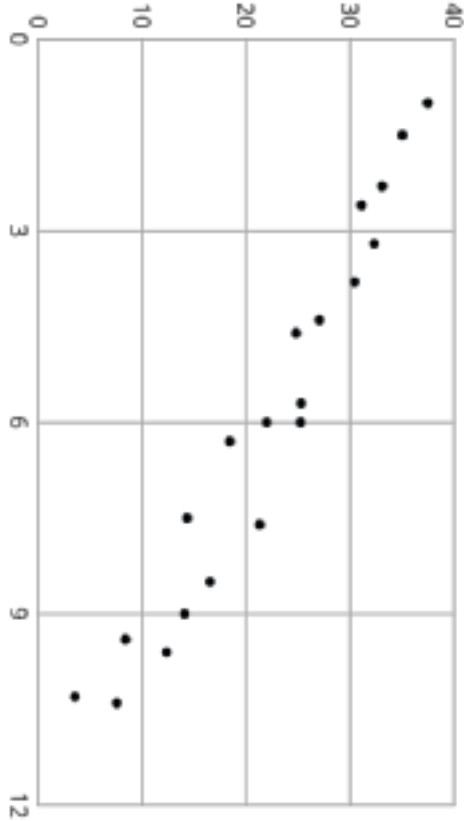
Scatter Plot City
C



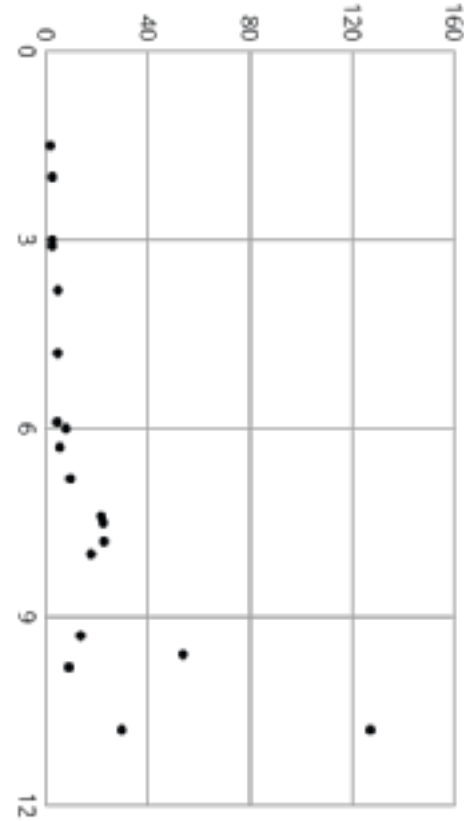
Scatter Plot City
D



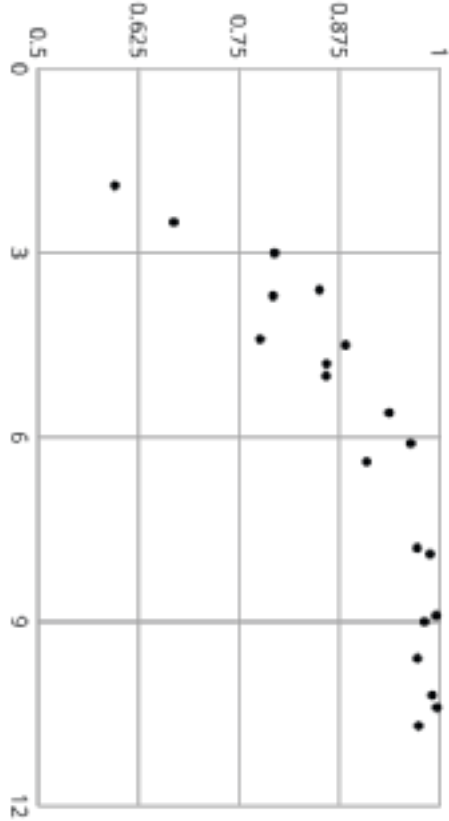
Scatter Plot City
E



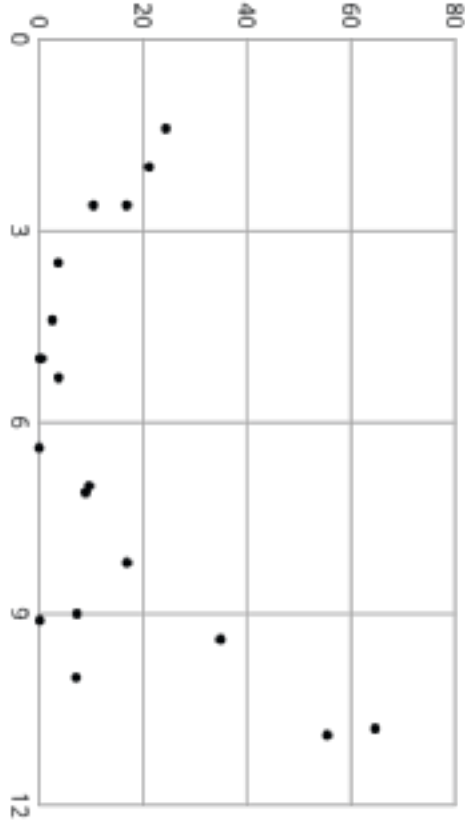
Scatter Plot City
F



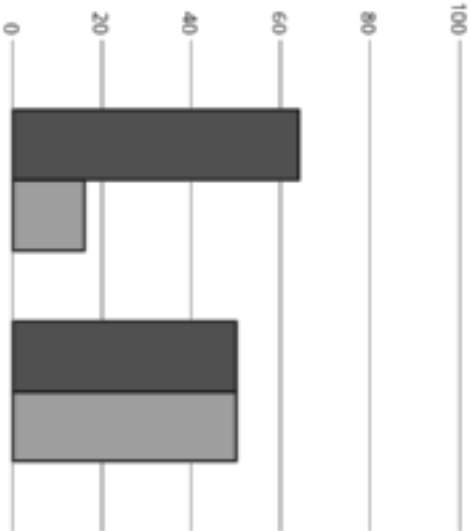
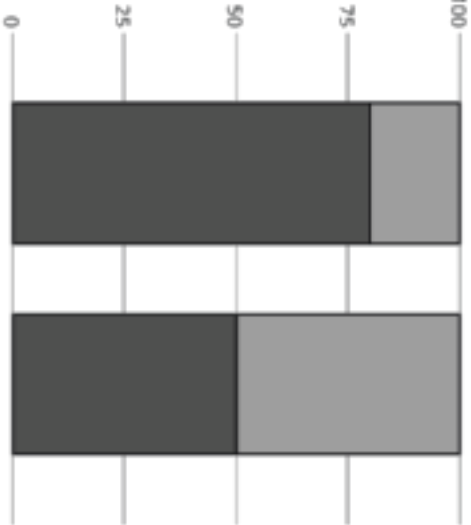
Scatter Plot City
G



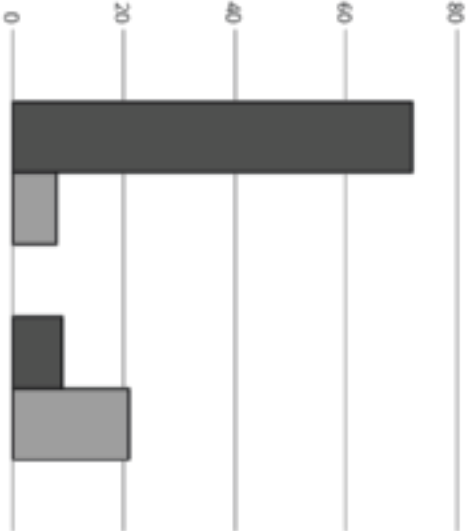
Scatter Plot City
H

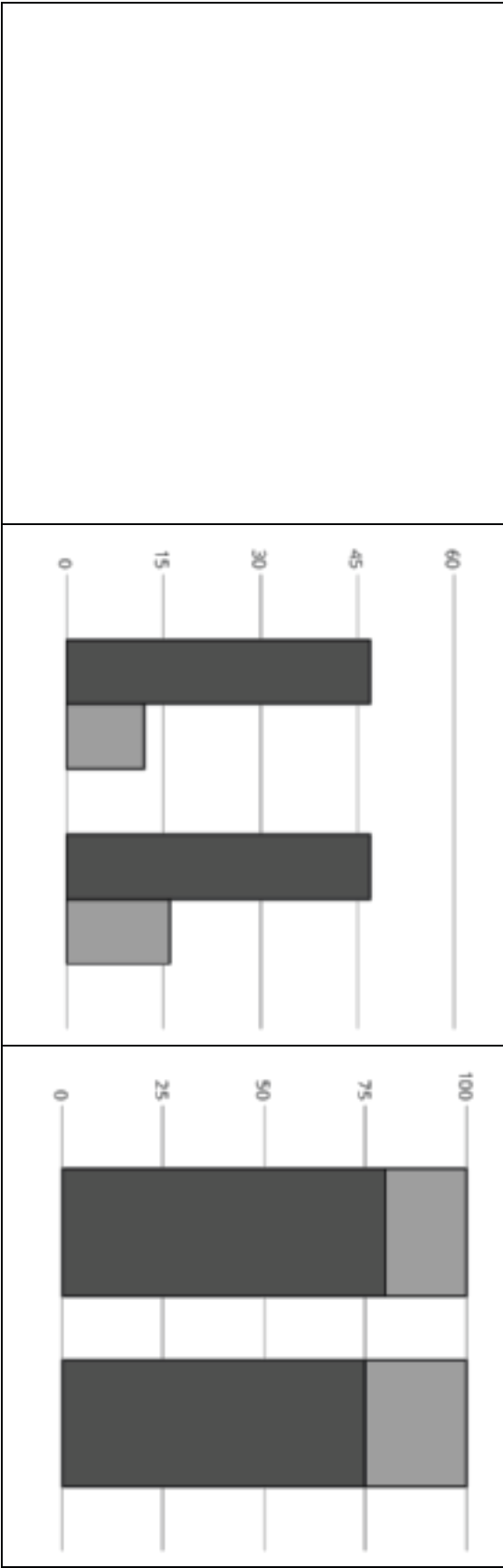


	liked the movie	disliked the movie	total
read the book	64	16	80
did not read the book	50	50	100
total	114	66	180



	satisfied with service	dissatisfied with service	total
satisfied with food	72	8	80
dissatisfied with food	9	21	30
total	81	29	110





Credits

CKMath K–8 was originally developed by Open Up Resources and authored by Illustrative Mathematics, <https://www.illustrativemathematics.org>, and is copyrighted as 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). The Open Up Resources K–8 Math Curriculum is available at: <https://www.openupresources.org/math-curriculum/>.

Adaptations and updates to the IM K–8 Math English language learner supports are copyright 2019 by Open Up Resources and licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0),

Adaptations and updates to IM K–8 Math are copyright 2019 by Illustrative Mathematics, including the additional English assessments marked as "B", and the Spanish translation of assessments marked as "B". These adaptations and updates are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

This particular work is based on additional work of the Core Knowledge® Foundation (www.coreknowledge.org) made available through licensing under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Illustration and Photo Credits

Ivan Pestic / Cover Illustrations

Illustrative Math K–8 / Cover Image, all interior illustrations, diagrams, and pictures / Copyright 2019 / Licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

These materials include public domain images or openly licensed images that are copyrighted by their respective owners, unless otherwise noted/credited. Openly licensed images remain under the terms of their respective licenses.



CKMath™
Core Knowledge **MATHEMATICS™**

CKMath™
Core Knowledge MATHEMATICS™

A comprehensive program for mathematical skills and concepts
as specified in the **Core Knowledge Sequence**
(content and skill guidelines for Grades K–8).

Core Knowledge MATHEMATICS™
units at this level include:

Rigid Transformations and Congruence
Dilations, Similarity, and Introducing Slope
Linear Relationships

Linear Equations and Linear Systems
Functions and Volume

Associations in Data

Exponents and Scientific Notation
Pythagorean Theorem and Irrational Numbers
Putting It All Together

www.coreknowledge.org

Core Knowledge Curriculum Series™

ISBN: 979-8-88970-847-6