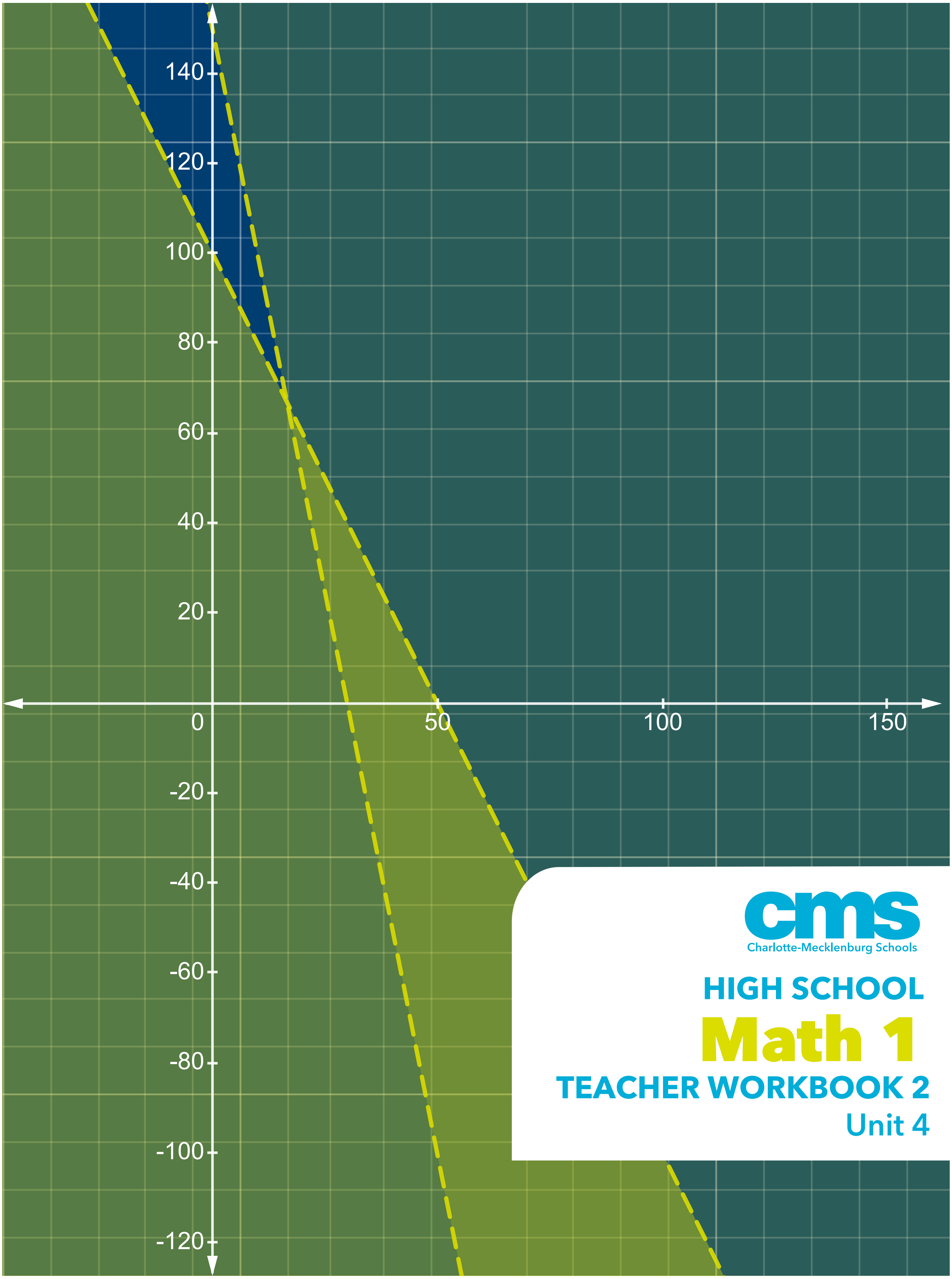


cms
Charlotte-Mecklenburg Schools

HIGH SCHOOL

Math 1

TEACHER WORKBOOK 2

Unit 4

Table of Contents

Unit 4: Two-variable Statistics

In grade 8, students informally constructed scatter plots and lines of fit and noticed linear patterns. In this unit, students build on this previous knowledge by assessing how well data can be represented by a linear model: first informally, then using the correlation coefficient, and then using residuals. Students learn to use technology to create, rather than estimate, a line of best fit.

Students begin Lesson 1 by reviewing statistical versus non-statistical questions. They then use data sets to help answer two of the statistical questions: median income in a zip code versus number of organic food items at grocery stores, and number of absences versus grade. Students discuss the real-world implications of these relationships and have opportunities to discuss the impacts of food deserts. At the same time, they learn that relationships between two variables have a form (linear or nonlinear) and a direction (positive, negative, or neither). Lesson 2 introduces the third characteristic of bivariate data: strength. At this point, students learn that the correlation coefficient, r, is a way to quantify the strength and direction of a linear relationship. They examine scatter plots showing different r-values and learn to get a feel for what the number represents.

Lessons 3 and 4 revisit lines of best fit, which are types of linear models. Students interpret the slope and y-intercept of a line of best fit in context and use the line of best fit to make predictions. They learn to use technology to determine the equation of the line of best fit (sometimes called the "regression equation") and the r-value for a data set. They also learn that linear models may not be useful far away from the data points that were used to create the best fit line, and they should be cautious and think carefully about whether the predicted values they find are reasonable. One data set represents COVID cases in NC in 2021.

Lessons 5 and 6 are Checkpoint Lessons. Most of the station activities have students apply their knowledge from this unit. In one station, they investigate characteristics of schools in their district or state and in another, they revisit the relationship between zip codes and availability of organic produce In addition, one station introduces students to calculating average rate of change, one revisits solving systems of equations from Unit 3, and another has students demonstrate a surprising result using coordinate geometry.

In Lesson 7, students learn about residuals, which measure how closely the line of best fit matches the actual data values. Students learn to use residual plots to assess the goodness of a linear model. In the Lesson 7 bridge, students examine the relationship between obesity and participation in SNAP (food stamps) in Baltimore, MD. And in Lesson 8, students are cautioned that even a strong association between variables does not necessarily mean that one causes the other. In an optional Lesson 8 activity (that could be used any time after Lesson 4), students make and evaluate a prediction about the height of an ancient human based on the length of their humerus.

Lessons 9 and 10 are Mathematical Modeling lessons. These lessons contain two new modeling prompts. In the first prompt, students choose a heating (and perhaps air conditioning) system based on efficiency and cost. In the second prompt, students explore relationships between characteristics of colleges: for instance, average cost of attendance versus average SAT scores. Prompts that were not used in Unit 2 may also be used here.

Lesson 11 occurs after administering the Unit 4 assessment and includes post-assessment activities. Taking this time to pause after the assessment to collect student reflection data through a survey and teacher conferences is a critical aspect of the course and building the classroom culture. Additionally, one activity supports students in adding and subtracting linear expressions in preparation for work in future units.

Instructional Routines

Aspects of Mathematical Modeling: Lessons 5 & 6, 8, 9 & 10

Card Sort: Lessons 1, 2, 7, 8

Co-Craft Questions (MLR5): Lesson 3

Collect and Display (MLR2): Lessons 1, 2, 7

Critique, Correct, Clarify (MLR3): Lesson 4

Discussion Supports (MLR8): Lessons 1, 2, 3, 7

Fit It: Lessons 4, 7

Math Talk: Lessons 2, 7

Notice and Wonder: Lessons 1, 3, 8, 11

Round Robin: Lessons 9 & 10

Stronger and Clearer Each Time (MLR1): Lessons 7, 8

Take Turns: Lessons 1, 2

Lesson 1: Describing Scatter Plots

PREPARATION

Lesson Goals	Learning Targets
• Describe (orally and in writing) the shape and direction of the relationship between two variables. • Use technology to create a scatter plot.	• I can describe the shape and direction of the relationship between two variables. • I can use technology to create a scatter plot to show the relationship between two variables.

Lesson Narrative

This lesson builds on students' work with bivariate data from grade 8 and their work with statistics in Unit 1. The purpose of this lesson is to describe the relationship between two variables in regards to the shape of the data (**linear** or **nonlinear**) and direction (**positive**, **negative**, or no relationship). The work in this lesson connects to Lesson 2, where students interpret the correlation coefficient as a measure of strength and direction of a linear relationship.

The lesson begins by revisiting statistical questions. Students identify the statistical questions and discuss what data they would need to collect to help study the question. Next, students create scatter plots, using technology, and use these to describe the relationship between two variables: household income and availability of organic vegetables. Lastly, students will examine a variety of scatter plots to distinguish between the different descriptions of shape and direction. Students will access data from a spreadsheet that allows them to copy and paste as needed; data sets are provided on different tabs of the spreadsheet aligned with the component of the lesson: http://bit.ly/U4L1DataSet.

When students analyze the scatter plot to describe the relationship between two variables, they are reasoning abstractly and quantitatively (MP2). The use of technology to create and analyze scatter plots is building the practice of using appropriate tools strategically (MP5). As students use appropriate and more specific vocabulary to describe the relationships, they are attending to precision (MP6).

What are you excited for your students to be able to do after this lesson?

Focus and Coherence

Building On	Addressing
NC.8.SP.1: Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Investigate and describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.	**NC.M1.S-ID.6:** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

Agenda, Materials, and Preparation

A data spreadsheet will be used throughout components of this lesson. Students can access this data through this link: http://bit.ly/U4L1DataSet.

- **Bridge** (*Optional, 5 minutes*)
- **Warm-up** (*10 minutes*)
- **Activity 1** (*10 minutes*)
  - Technology is required for this activity: Acquire devices that can access Desmos (recommended) or other graphing technology. It is ideal if each student has their own device.
  - Food Desert Statistics handout (print 1 copy per student)
- **Activity 2** (*10 minutes*)
  - Shape and Direction card sort (print 1 copy per group of students and cut up in advance)
- **Lesson Debrief** (*5 minutes*)
- **Cool-down** (*5 minutes*)
  - M1.U4.L1 Cool-down (print 1 copy per student)

LESSON

Bridge (*Optional, 5 minutes*)

Addressing: NC.8.SP.1

In this bridge, students practice identifying and interpreting points in a scatter plot. Encourage students to think about how the scatter plot is constructed. This task is aligned to question 1 in Check Your Readiness.

RESPONSIVE STRATEGIES

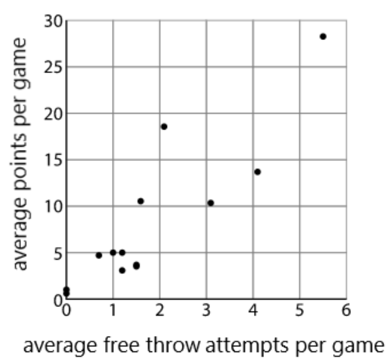
Provide an alternate table with visual cues for each column. For example, a picture of a basketball player under "player," a basketball and goal under "average free throw attempts," and a score board under "average points per game."

Encourage students highlight the "average free throw attempts" in the table and the corresponding axis the same color and "average points per game" and the corresponding axis a different color.

Student Task Statement

Here is a table and a scatter plot that compares points per game to free throw attempts for a basketball team during a week-long tournament.[1]

1. Circle the point that represents the data for Player E.
2. What does the point $(2.1, 18.6)$ represent?
3. 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.



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

[1] Adapted from IM 6–8 Math https://curriculum.illustrativemathematics.org/MS/index.html, which was originally developed by Open Up Resources and authored by Illustrative Mathematics, and is copyright 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). OUR's 6–8 Math Curriculum is available at https://openupresources.org/math-curriculum/. Adaptations and updates to IM 6–8 Math are copyright 2019 by Illustrative Mathematics, and are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

DO THE MATH	PLANNING NOTES

Warm-up: Statistical Questions (*10 minutes*)

Instructional Routines: Notice and Wonder; Discussion Supports (MLR8) - Responsive Strategy	
Building On: NC.8.SP.1	**Addressing:** NC.M1.S-ID.6

In Unit 1, students investigated univariate data and created data displays such as histograms and box plots. In this warm-up, students revisit statistical questions and discuss the data needed to investigate the question.

Next, the discussion focuses on the second part of the question: "How does daily attendance in school impact student achievement?" Students learn to create scatter plots using Desmos or other graphing technology using data on the number of absences in a course and grade in a course.

The relationship between these variables has a negative direction, meaning that as one variable increases the other variable decreases. This scatter plot will be used later, along with the scatter plot in Activity 1, to introduce vocabulary for direction (positive or negative). The focus in this warm-up is to recognize when a statistical question may involve two variables and that scatter plots are tools for displaying the relationship between the two variables.

Step 1

- Ask students to arrange themselves into pairs or use visibly random grouping.
- Provide students with 2 minutes of individual quiet think time to work on questions 1 and 2.
- Ask students to share with their partner their response to question 2.

Student Task Statement

1. Select **all** of the statistical questions.
    a. What is the current value of a 2,000-square-foot house in Charlotte, NC?
    b. How does daily attendance in school impact student achievement?
    c. How many counties are there in North Carolina?
    d. Does the neighborhood someone lives in determine their access to healthy foods?
2. Choose one of the statistical questions and describe what data would need to be collected to help answer the question.

Step 2

- Invite students to share the question shared by their partner and what data could be used to investigate the question.
- Use student responses to distinguish between univariate data such as the prices of 2,000-square-foot homes and bivariate data such as attendance and grades.
- Record the suggested data for question 1d for all to see. In Activity 1, these suggestions will be used to introduce the data provided to investigate this question.

RESPONSIVE STRATEGY

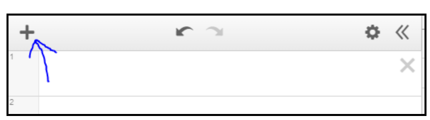
As students are sharing responses from question 2, display sentence frames for students to use , such as: "The data the needs to be collected to answer this question is . . .," "I chose the question_____," and "To answer this question, I will need to collect the following data . . . ."

Discussion Supports (MLR8)

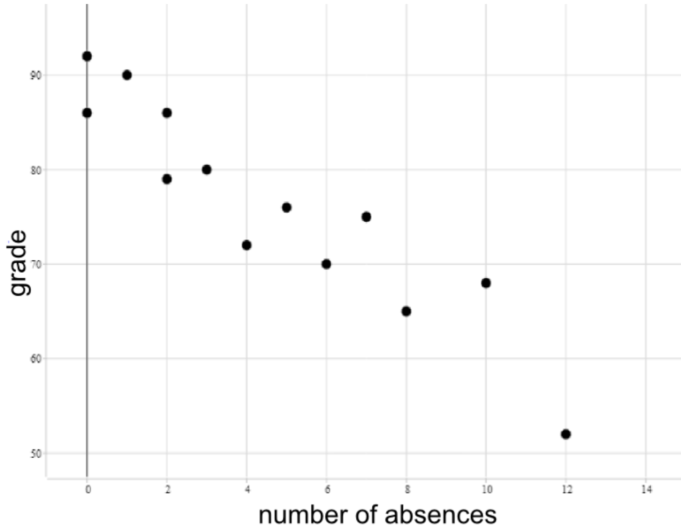
Absences	Grade
0	86
0	92
1	90
2	86
2	79
3	80
4	72
5	76
6	70
7	75
8	65
10	68
12	52

Step 3

- Data, such as the number of absences and the grade in a class, can help to investigate how daily attendance in school impacts student achievement. Display the data from the Warm-up tab (http://bit.ly/U4L1DataSet) and ask students what they notice and wonder using the *Notice and Wonder* routine.
  - Provide students with 1 minute of quiet think time and then have students turn and share with their partner.
  - Invite students to share. Record and display their responses for all to see.
- Facilitate a discussion on the trend in the data and how to display the data. Possible questions to include are:
  - If not mentioned during Notice and Wonder, ask students: “How do the grades change as the number of absences increase?” (The grades decrease.)
  - “The data are presented in a table. What other ways can we display this data?” (scatter plot)
  - “What would the horizontal axis represent? What would the vertical axis represent?” (the number of absences, the grade in a class)
  - “What would a point represent?” (the data about one student: number of absences and grade in the class)
- Demonstrate how to use Desmos graphing calculator to create a scatter plot.
  - Click the + icon in the top left corner of the entry panel, select table, and begin entering the data.



  - As you enter, coordinate pairs will be plotted on the graph. If needed, adjust the graphing window to see the plotted points. Use the magnifying glass located below and to the left of the table to “zoom fit” the graph settings to the data.



- Data sets can be copied from other sources such as a spreadsheet and pasted into Desmos. This is helpful when dealing with large data sets. Demonstrate how to copy and paste the data.

- Ask students:
  - "What relationship is indicated in the scatter plot?" (The more absences, the lower the grade.)
  - "How is that shown in the scatter plot?" (When reading the graph from left to right, the points are going down.)
- Tell students they will now analyze data that will help investigate question 1d from the warm-up.

DO THE MATH	PLANNING NOTES

Activity 1: Income and Food Access[2] (*10 minutes*)

Instructional Routine: Collect and Display (MLR2)	
Building On: NC.8.SP.1	**Addressing:** NC.M1.S-ID.6

In this activity, students explore the question: "Does your neighborhood determine access to healthy foods?" The data set provided was collected in 2019 in grocery stores within San Antonio, TX and includes the average household income of the neighborhood and the number of organic vegetables offered at local stores. Neighborhoods were identified by zip code.

Students create a scatter plot and describe the relationship between the two variables. The data in this activity has a positive direction, meaning that as one variable increases, the other increases.

[2] Adapted from Skewthescript.org

Step 1

- Keep students in pairs.
- Write the term "food desert" on the board.
- Ask the class: "What do you think the term 'food desert' means?" (A food desert is an area -neighborhood, community, etc.- where healthy, affordable food is difficult to obtain.)
- Distribute the Food Desert Statistics handout[3] and have students read it.
- Ask pairs to choose one of the following questions[4] to answer and then report their answer to the class:
  - Why might healthy, affordable food be difficult to obtain in certain areas?
  - In which types of areas/communities do you think food deserts are most prevalent: urban, rural or suburban? Do they only exist in those areas?
  - How do you think living in a food desert could affect a person's/family's food choices?
- Revisit the data students suggested during the warm-up to address the question about access to healthy foods. Introduce the data provided on average household income and number of organic vegetables offered. Ask students:
  - "How might these data help us investigate this question?" (There needs to be a number to describe the neighborhood. This could be the average household income. Another possibility could be the population. There also needs to be a number to pair with the neighborhood that describes healthy foods. This could be organic vegetables. It could also be square footage of the produce section in the local grocery store.")
  - "What concerns might you have about using these data?" (Vegetables do not have to be organic to be healthy. This narrow definition of "healthy foods" should be considered when making any conclusions from the analysis of the data.)

Step 2

- Provide students with a computer and access to Desmos. The Activity 1 data set is provided in the spreadsheet referenced in the Lesson Narrative to allow students to copy and paste the data.
- Provide students with a couple of minutes to individually create the scatter plot and then ask that they work with their partner to create a description of the relationship.
- Using the *Collect and Display* routine, listen for and scribe the language students use to describe the relationship between the two variables as they work. Be prepared to display students' words and phrases in Step 3.

> **Monitoring Tip:** Identify descriptions that vary in precision from the general to the more specific.
> - It goes up.
> - It goes up when the income goes up.
> - As the income increases, the number of organic vegetables offered increases.
> - The number of organic vegetables offered increases as the income increases. Once over $100,000, the number is mostly constant at 95 items.

[3] From Food Deserts: Causes, Consequences and Solutions. Reprinted with permission of Teaching Tolerance, a project of the Southern Poverty Law Center. https://www.learningforjustice.org/classroom-resources/lessons/food-deserts-causes-consequences-and-solutions
[4] From Food Deserts: Causes, Consequences and Solutions. Reprinted with permission of Teaching Tolerance (see above).

Student Task Statement

"Food security exists when all people, at all times, have physical and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life" (World Food Summit, 1996).

Students at a high school in San Antonio, TX decided to explore the access to healthy food items available at grocery stores in different neighborhoods. They decided to collect data on the average household income for the neighborhood (defined by zip code) and the number of organic items available in the local grocery store.

Average household income	Number of organic vegetables offered
$71,186	36
$34,234	4
$71,186	28
$48,760	31
$78,096	78
$40,506	14
$38,166	12
$50,398	18
$49,437	38
$66,073	84
$86,566	61
$78,176	56
$59,154	62
$50,252	44
$48,364	26
$56,274	29
$41,318	15
$125,145	95
$65,911	18

Average household income	Number of organic vegetables offered
$50,252	65
$53,945	50
$59,072	35
$49,437	36
$72,080	28
$108,486	95
$70,530	46
$57,199	29
$78,176	73
$78,288	53
$86,566	86
$84,181	68
$84,181	56
$78,176	85
$84,181	86
$135,547	93
$92,946	82
$77,894	96

1. Create a scatter plot to display the (*average household income, number of organic vegetables offered*).
    a. In www.desmos.com/calculator, click on the + icon on the top left of the window. Select the table option.
    b. Copy the data set provided at bit.ly/U4L1DataSet and paste into the first entry line in Desmos.
    c. As you enter, coordinate pairs will be plotted on the graph. If needed, adjust the graphing window to see the plotted points. Use the magnifying glass located below and to the left of the table to "zoom fit" the graph settings to the data.
2. Describe the relationship between the (*average household income, number of organic vegetables offered*).
3. Why are these data important to understand? Do you think a similar trend exists in other cities? Why or why not?

Are You Ready For More?

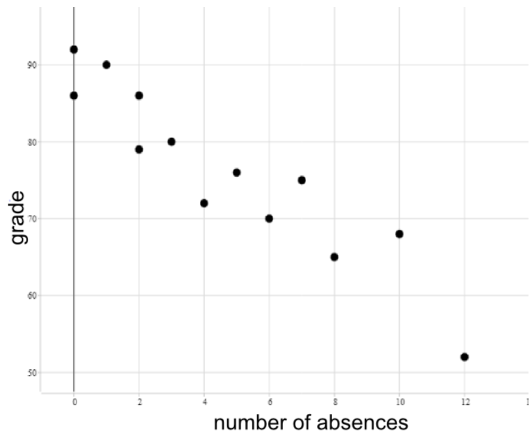
Students in Charlotte, NC were interested in examining the access in their city. They collected the following data. In this case, they also collected the population within the neighborhood (defined by zip code).

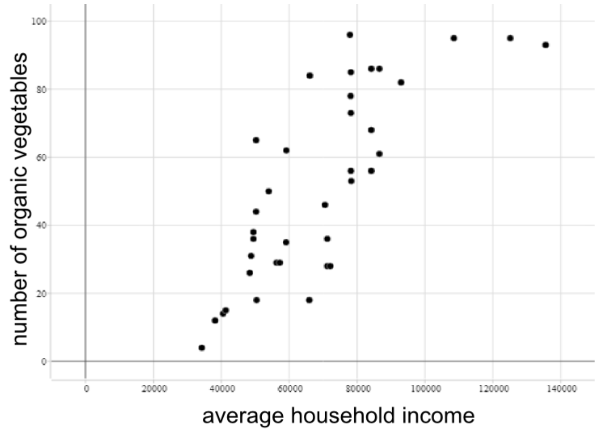
1. Create a scatter plot for the (*median household income, organic produce available*) and describe any relationship between the two variables.
2. Compare this relationship to the one you found for San Antonio. What do you think are the reasons for any similarities or differences?
3. Create a scatter plot for the (*population, organic produce available*) and describe any relationship between the two variables.
4. One of the points appears to be an outlier. How does your answer to question 3 change if the outlier is removed?

Population	Median household income (2019)	Organic produce available
71048	65963	27
59664	93942	40
49635	59438	43
9280	136333	44
53629	51676	44
37286	91494	44
37309	45808	46
11315	88039	47
11195	92786	55
43931	52766	55
42263	71914	55
19283	93938	56
28523	90057	57
20317	76022	58
47208	49465	59

Step 3

- Display the scatter plot for all to see.
- Select students to briefly share their responses.
- To help students see the connection between their descriptions, refer to a display of any student words and phrases collected as they worked. Use arrows or annotations to highlight connections between specific descriptions such as how “up” connects to “increases” and “it” refers to average household income or number of organic vegetables.
- Facilitate a discussion on the direction of data as displayed in a scatter plot.
  – Display the scatter plot from the warm-up on the *absences* and *grade* alongside the scatter plot for *average household income* and *number of organic vegetables*.





  – Ask students: “What is the direction of the relationship in each scatter plot?” (up, down, increasing, and/or decreasing)
    - What does this relationship indicate about food deserts in San Antonio, Texas?
  – Connect scatter plots, student language, and descriptions with the vocabulary of:
    - positive direction: as one variable increases, the other variable increases
    - negative direction: as one variable increases, the other variable decreases

DO THE MATH

PLANNING NOTES

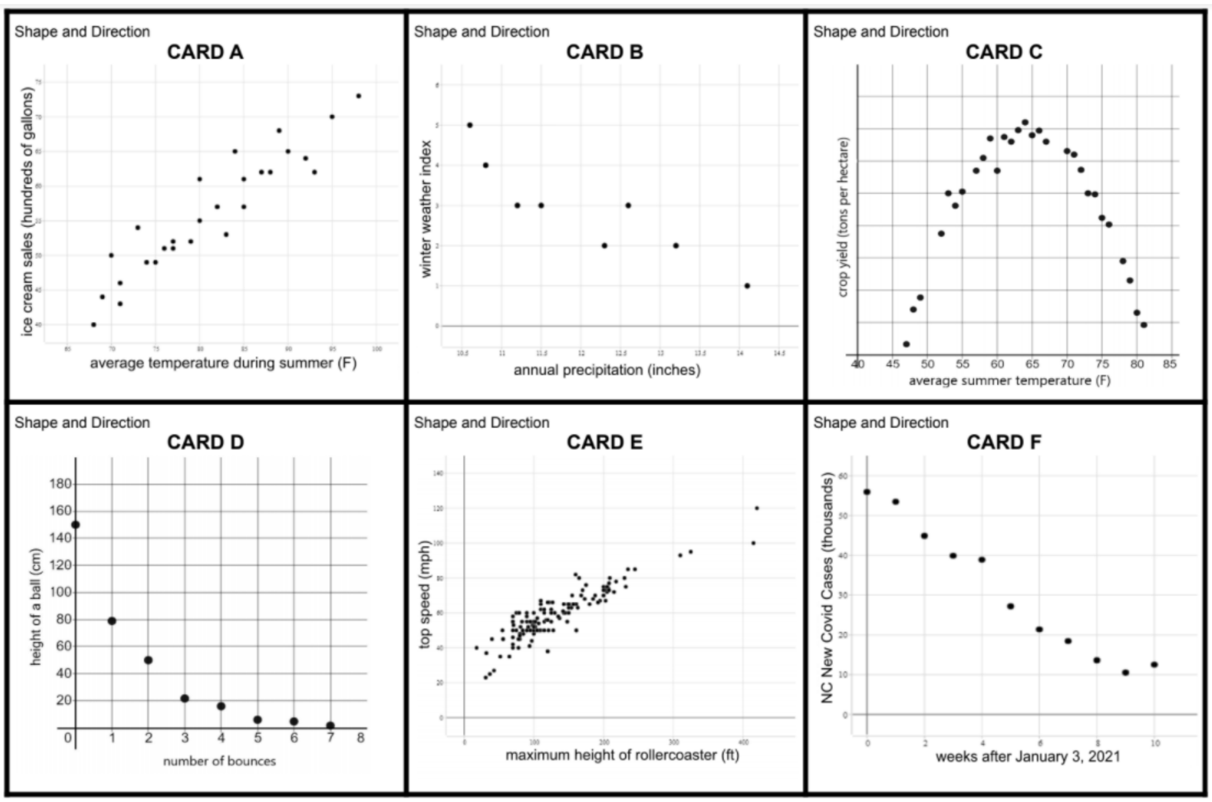
Activity 2: Shape and Direction (*10 minutes*)

Instructional Routines: Card Sort; Take Turns; Discussion Supports (MLR8) - Responsive Strategy	
Building On: NC.8.SP.1	**Addressing:** NC.M1.S-ID.6

In this *Card Sort* activity, students analyze a scatter plot to identify the shape (linear or nonlinear) and direction (positive, negative, no relationship).

As students use the *Take Turns* routine to explain their thinking to a partner, encourage them to use precise language and mathematical terms to refine their explanations (MP6).

Here is an image of the cards for reference and planning.



Step 1

- Keep students in pairs.
- Tell students they will be sorting a set of cards first by the shape, then by the direction of the data as shown in the scatter plot.

RESPONSIVE STRATEGY

As students take turns describing the relationship between two variables to their partner, display the following sentence frame for all to see: "As the _____ increases the _____ increases/decreases." For question 3, display "There appears to be a _____ _____ relationship between ____ and ____," and "There does not appear to be a relationship between _____ and ______." Encourage students to press each other for justifications when they disagree.

Discussion Supports (MLR8)

Student Task Statement

Your teacher will give you a set of cards, each with a different scatter plot. With a partner, complete each of the following:

1. Sort the cards based on the direction of the data displayed in the scatter plot. Record the card(s) for each group.

   Positive direction: ________________ Negative direction:________________ No Relationship: ________________

2. Sort the cards based on the shape of the data. Does the scatter plot have a linear or a nonlinear shape? Record the card(s) for each group.

   Linear shape: __________________ Nonlinear shape: _____________________

3. Take turns with your partner. Select one of the cards from the scatter plots. Describe the relationship between the two variables to your partner. For each description your partner shares, listen carefully to their explanation. If you disagree, discuss your thinking and work together to reach an agreement.

Step 2

- Ask pairs of students to volunteer to share the description for one of the scatter plots. After a pair shares, ask if other pairs agree or disagree. If there is disagreement, ask pairs to share their thinking and work towards an agreement.
- Attend to the language that students use in their descriptions by giving them opportunities to include information such as the specific variables.

DO THE MATH	PLANNING NOTES

Lesson Debrief (*5 minutes*)

The purpose of this lesson is for students to create scatter plots using technology and describe the relationship between the two variables with a focus on shape and direction.

Choose what questions to focus the discussion on, whether students should first have an opportunity to reflect in their workbooks or talk through these with a partner, and what questions will be prioritized in the full class discussion. Consider providing students with whiteboards to share their sketches.

- "How does a scatter plot help in analyzing the relationship between two variables?" (The scatter plot makes it easier to see any possible relationship between the two variables.)
- Ask students to sketch a scatter plot that has either a positive direction, negative direction, or no direction. Ask them to show their scatter plot to the class and have students determine the direction.
- "What information does the direction provide about the relationship between two variables?" (The direction indicates that as one variable increases, the other variable increases if positive or decreases if negative.)
- Ask students to sketch two scatter plots: one with a linear shape and the other with a nonlinear shape. Ask them to label the scatter plots with the direction (positive, negative, or neither).

PLANNING NOTES

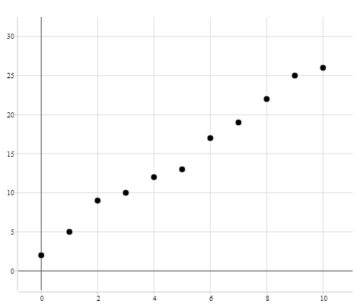
Student Lesson Summary and Glossary

A scatter plot is used to display two numerical variables. Scatter plots are helpful when determining the relationship between the two variables. This relationship can be described in terms of the shape and the direction.

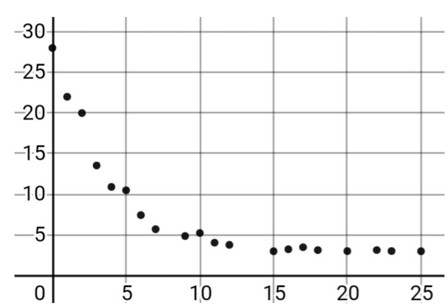
The shape of data in a scatter plot can be described as **linear** or **nonlinear**.

Linear relationship: A relationship between two numerical variables in which the scatter plot displaying the data resembles a line.

Nonlinear relationship: A relationship between two numerical variables in which the scatter plot displaying the data has a pattern other than a line, usually curved in some way.

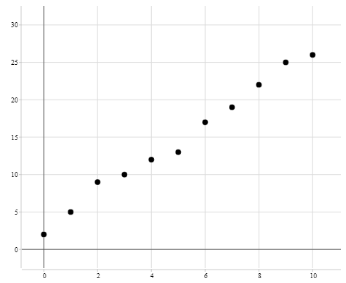


The data points in this scatter plot show a **linear** relationship. The data closely resemble a straight line.

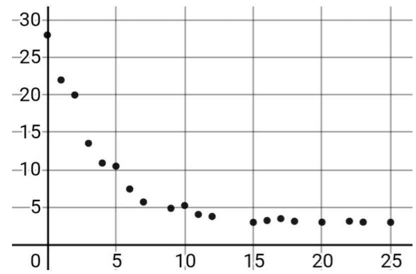


The data points in this scatter plot show a **nonlinear** relationship. The data do not resemble a straight line.

The direction of the data in a scatter plot can be described as **positive** or **negative**.



The data points in this scatter plot show a **positive** relationship. As the x variable increases, the y variable increases.



The data points in this scatter plot show a **negative** relationship. As the x variable increases, the y variable decreases.

Positive relationship: A relationship between two numerical variables in which one variable tends to be paired with an increase in the other variable.

Negative relationship: A relationship between two numerical variables in which one variable tends to be paired with a decrease in the other variable.

The shape and direction of the data as shown in the scatter plot help when describing the relationship.

The scatter plot on the left displays the relationship between blood alcohol level and reaction time. There is a **positive, linear** relationship. This means that as the blood alcohol level increases, the reaction time increases.

Cool-down: Clear Skies (*5 minutes*)

Addressing: NC.M1.S-ID.6

Cool-down Guidance: More Chances
In the next lesson, students will interpret the correlation coefficient, which will inform them of the strength and direction of a linear relationship. They will have additional opportunities to describe the relationship between two variables.

Cool-down

The National Climate Data Center collects data on weather conditions at various locations. The table below shows data for elevation above sea level and the mean number of clear days per year for 14 U.S. cities.[5]

1. Copy the data set provided on the right from bit.ly/U4L1DataSet on the Cool-down tab. Use technology to create a scatter plot of the *(elevation above sea level, mean number of clear days)* data.

2. Identify the shape and direction of the relationship.

3. Describe the relationship between elevation above sea level and the mean number of clear days.

City	Elevation above sea level (feet)	Mean number of clear days per year
Albany, NY	275	69
Albuquerque, NM	5,311	167
Anchorage, AK	114	40
Boise, ID	2,838	120
Boston, MA	15	98
Helena, MT	3,828	82
Lander, WY	5,557	114
Milwaulkee, WI	672	90
New Orleans, LA	4	101
Raleigh, NC	434	111
Rapid City, SD	3,162	111
Salt Lake City, UT	4,221	125
Spokane, WA	2,356	86
Tampa, FL	19	101

Student Reflection: What new information did you learn about food security and neighborhoods during this lesson? What additional information are you interested in understanding around this topic? What could you do to help change the reality of food insecurity?

[5] Adapted from EngageNY https://www.engageny.org/ for the New York State Department of Education, which was originally developed and authored by Great Minds. It is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 United States (CC BY-NC-SA 3.0 US).

DO THE MATH

INDIVIDUAL STUDENT DATA

SUMMARY DATA

NEXT STEPS

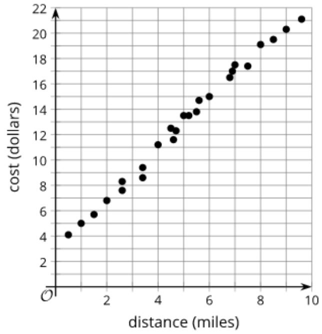
TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

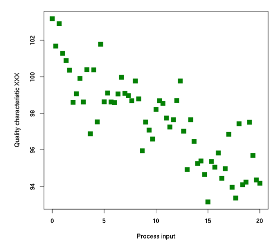
Based on the new information students expressed they learned in this lesson in their student reflection, what are some ways you can support their real-world efforts for change?

Practice Problems

1. The scatter plot shows the relationship between *(distance, cost)*. Describe the shape and direction of the relationship.



2. Here is a scatter plot:



Select **all** of the following that describe the relationship in the scatter plot:

   a. Linear relationship
   b. Non-linear relationship
   c. Positive relationship
   d. Negative relationship
   e. No relationship

3. (*Technology required*.) The table includes data collected on popcorn sales at a local carnival.

Popcorn price (dollars)	0.50	0.60	0.8	1	1.25	1.50	1.75
Average number of cups sold	150	136	125	120	105	80	75

   a. Use technology to create a scatter plot of the data.

   b. Describe the relationship between the price of the popcorn and the average number of cups sold.

4. (*Technology required*.) The Census Bureau provided the following statistics for the years from 2006–2011.[6]

Year	Median income for all men
2011	$37,653
2010	$38,014
2009	$38,558
2008	$49,134
2007	$41,033
2006	$41,103

Year	Median income for all women
2011	$23,395
2010	$23,657
2009	$24,284
2008	$23,967
2007	$25,005
2006	$24,429

   a. Use technology to create one scatter plot for the median income for all men from 2006–2011 and one scatter plot for the median income for all women from 2006–2011. Describe any relationship between the year and the income for each scatter plot.

   b. Compare this relationship of median incomes for men and women. What do you think are the reasons for any similarities or differences?

[6] Adapted from Secondary Math 1, Module 9, Mathematics Vision project http://www.mathematicsvisionproject.org, licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0)

5. (*Technology required*.) Different stores across the country sell the same book for different prices. The table shows the price of a particular book in dollars and the number of books sold at that price.[7]

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

a. Create a scatter plot using technology

b. Identify the shape and direction of the relationship.

c. Describe the relationship between price charged for the book and the number sold.

d. Is there an outlier? What point? How is the relationship affected if the outlier is removed?

6. The robotics team needs to purchase \$350 in new equipment. Each student on the team plans to fundraise and contribute equally to the purchase. If X represents the total number of students on the team, which expression represents the amount each student needs to contribute? Which expression represents the amount each student needs to fundraise?

a. $\$350 - X$

b. $\$350 + X$

c. $\frac{\$350}{X}$

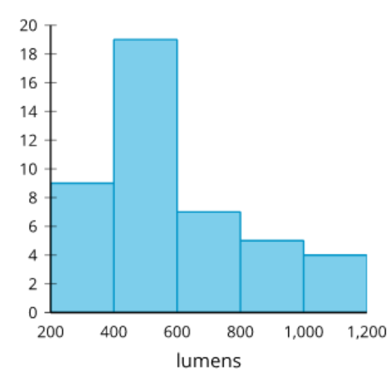
d. $\$350 \cdot X$

Amount each student needs to contribute	Amount each student needs to fundraise

(From Unit 2)

7. Describe the shape of the distribution shown in the histogram that displays the light output, in lumens, of various light sources.

(From Unit 1)



[7] Adapted from IM 6–8 Math https://curriculum.illustrativemathematics.org/MS/index.html, which was originally developed by Open Up Resources and authored by Illustrative Mathematics, and is copyright 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). OUR's 6–8 Math Curriculum is available at https://openupresources.org/math-curriculum/. Adaptations and updates to IM 6–8 Math are copyright 2019 by Illustrative Mathematics, and are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

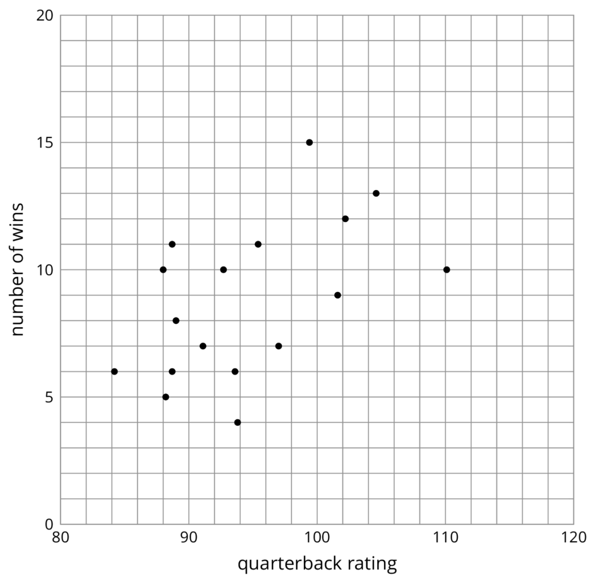
8. Here is a table and scatter plot that show ratings and wins for quarterbacks who started 16 games this season.[8]

   a. Circle the point in the scatter plot that represents Player K's data.

   b. What does the point $(88, 10)$ represent?

   c. 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 eight games. On the scatter plot, plot a point that represents Player R's data.

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

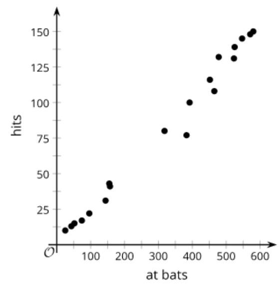


(Addressing NC.8.SP.1)

9. The scatter plot shows the number of times a player came to bat and the number of hits they had. The scatter plot includes a point at $(318, 80)$.

   Describe the meaning of this point in this situation.

   (Addressing NC.8.SP.1)



[8] Adapted from IM 6–8 Math https://curriculum.illustrativemathematics.org/MS/index.html, which was originally developed by Open Up Resources and authored by Illustrative Mathematics, and is copyright 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). OUR's 6–8 Math Curriculum is available at https://openupresources.org/math-curriculum/. Adaptations and updates to IM 6–8 Math are copyright 2019 by Illustrative Mathematics, and are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Lesson 2: The Correlation Coefficient

PREPARATION

Lesson Goal	Learning Target
• Describe (orally and in writing) the strength and sign of the relationship between variables based on the correlation coefficient.	• I can describe the strength of a relationship between two variables.

Lesson Narrative

The mathematical purpose of this lesson is for students to interpret the correlation coefficient for a bivariate, numerical data set. The work of this lesson builds from the prior lesson where students analyzed scatter plots to describe the direction and shape of a relationship. The **correlation coefficient** is introduced as a statistic that can be used to describe the strength and direction of a linear relationship. It serves to quantify how close to linear a data set is. Let students know that in Math 1, they'll be using technology to calculate the correlation coefficient. This will be addressed in Lesson 4. The focus in this lesson is on understanding how to interpret the correlation coefficient. If students inquire, share the Pearson's correlation coefficient formula and let students know that this is something they would use in a Statistics course:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

When students use the value of the correlation coefficient to describe the relationship between two variables, they are looking for and making use of structure (MP7). To make sense of the relationship between variables, students reason abstractly and quantitatively (MP2). When students examine relationships to think about correlations, they also consider additional variables that might have an influence on any trends they see. Deciding which variables need to be included is a part of the process of modeling with mathematics (MP4).

What math language will you want to support your students with in this lesson? How will you do that?

Focus and Coherence

Building On	Addressing
NC.5.NBT.3 Read, write, and compare decimals to thousandths. • Write decimals using base-ten numerals, number names, and expanded form. • Compare two decimals to thousandths based on the value of the digits in each place, using >, =, and < symbols to record the results of comparisons.	**NC.M1.S-ID.6:** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related. **NC.M1.S-ID.8:** Analyze patterns and describe relationships between two variables in context. Using technology, determine the correlation coefficient of bivariate data and interpret it as a measure of the strength and direction of a linear relationship. Use a scatter plot, correlation coefficient, and a residual plot to determine the appropriateness of using a linear function to model a relationship between two variables.

Agenda, Materials, and Preparation

- **Bridge** (*Optional, 5 minutes*)
- **Warm-up** (*5 minutes*)
- **Activity 1** (*15 minutes*)
  - Scatter Plot Fit card sort (print 1 copy per group of students and cut up in advance)
- **Activity 2** (*Optional, 10 minutes*)
- **Activity 3** (*10 minutes*)
  - Chart paper and markers
- **Lesson Debrief** (*5 minutes*)
- **Cool-down** (*5 minutes*)
  - M1.U4.L2 Cool-down (print 1 copy per student)
- **Practice Problems**
  - Data Spreadsheet for #7: https://bit.ly/U4L2DataSet

LESSON

Bridge (*Optional, 5 minutes*)

Instructional Routine: Math Talk	
Building On: NC.5.NBT.3	

The purpose of this *Math Talk* is to elicit strategies and understandings students have for ordering positive and negative decimals. These understandings help students develop fluency and prepare students to interpret correlation coefficients.

Student Task Statement

Mentally order the numbers from least to greatest.

a. 20.2, 18.2, 19.2

b. -14.6, -16.7, -15.1

c. -0.43, -0.87, -0.66

d. 0.50, -0.52, 0.05

DO THE MATH

PLANNING NOTES

Warm-up: Putting the Numbers in Context (*5 minutes*)

Addressing: NC.M1.S-ID.6

The mathematical purpose of this activity is for students to match a bivariate data set with its context. Students should think about whether they might expect a strong relationship or not as well as whether the relationship has a positive or negative relationship. Monitor for students who discuss linear relationships or variability in the data.

Step 1

- Ask students to arrange themselves in groups of two to four or use visibly random groupings.
- Provide access to the scatter plots and contexts for the warm-up. As students work, ask them to think about whether they would expect a strong relationship or a weaker relationship between any of the pairs of variables.

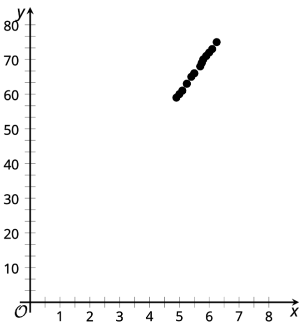
Advancing Student Thinking: Students may struggle with matching the pairs of variables with a scatter plot. Encourage students to think about how related the variables are and how the y variable might change as the x variable increases.

Student Task Statement

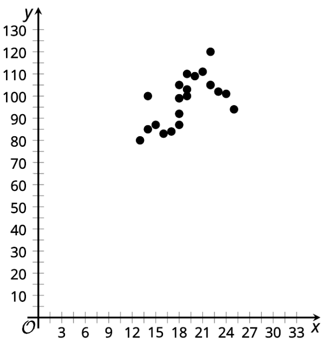
Match the variables to the scatter plot you think they best fit. Be prepared to explain your reasoning.

	x **variable**	y **variable**
1.	daily low temperature in Celsius for Denver, CO	boxes of cereal in stock at a grocery in Miami, FL
2.	average number of free throws shot in a season	basketball team score per game
3.	measured student height in feet	measured student height in inches
4.	average number of minutes spent in a waiting room	hospital satisfaction rating

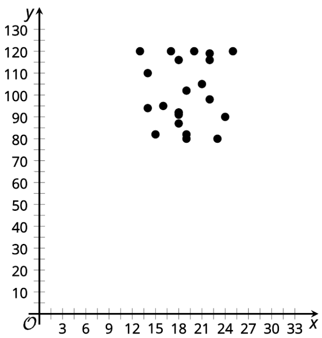
a.



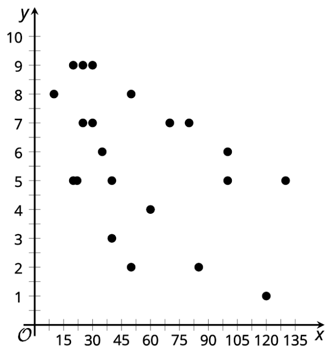
b.



c.



d.



Step 2

- Facilitate a whole-class discussion. The goal of this discussion is for students to discuss how the characteristics of the scatter plots allowed them to determine the context. For each context, select groups to share their match and reasoning. Select groups who used the direction and the shape in their small-group discussion.
  - "How did the shape help you make a match?" (Scatter plot C has a nonlinear shape. I expected the temperature and cereal context relationship to be totally random rather than linear.)
  - "How did the direction help you make a match?" (Scatter plot D has a negative direction. I assumed that the more time a person waited in a waiting room, the lower their satisfaction rating.)
  - "How did shape and direction help you make a match?" (Both scatter plots A and B have a positive direction. For the height in inches and the height in feet, I knew that the data would be closer to linear so it would match scatter plot A.)

DO THE MATH	PLANNING NOTES

Activity 1: Scatter Plot Fit (*15 minutes*)

Instructional Routine: Card Sort
Addressing: NC.M1.S-ID.8

In this *Card Sort* activity, students are given cards displaying scatter plots that have varying strengths and directions. Building from the previous lesson, students will first sort the scatter plots by direction (positive, negative and neither). Next, students order the scatter plots based on the strength of a linear shape.

The diagram intentionally resembles a number line, which will be used to introduce the correlation coefficient during the whole-class discussion. Consider providing groups with chart paper to draw the diagram and place the cards in order. Groups can label the correlation coefficient along the number line during discussion.

If time allows, start the activity by having groups sort the cards first based on their own criteria. If the criteria provided in the task emerge, build from there to introduce the correlation coefficient.

Step 1

- Ask students to arrange themselves into pairs or use visibly random groupings.
- Tell students that in this activity they will sort cards based on direction and strength of a linear relationship.
- Distribute one copy of the blackline master to each group.

Student Task Statement

Your teacher will give you a set of cards that show scatter plots of data.

1. Sort the scatter plots by direction.

Scatter plots with a negative direction	Scatter plots without direction	Scatter plots with a positive direction

2. Using the diagram below, organize the scatter plots based on the strength of the linear shape.

strongest linear shape negative direction — *weakest linear shape without direction* — *strongest linear shape positive direction*

Step 2

- Invite groups of students to share their ordering of the scatter plots. Display for all to see.
- Introduce the correlation coefficient as a statistical measure that can be used to interpret the strength and direction of a linear relationship.
- Display the correlation coefficients alongside the scatter plots. Tell students that the letter "r" is often used to report the correlation coefficient.

D	C	A	H	B	E	F	G
$r = -1$	$r = -0.88$	$r = -0.61$	$r = -0.13$	$r = -0.06$	$r = 0.46$	$r = 0.94$	$r = 1$

- Provide students with 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 in which students share what they noticed.
- Among things students should notice are:
  - The sign is the same as the direction.
  - The values of r seem to go from -1 to 1.
  - The closer r is to 1 or -1, the stronger the linear relationship between the variables.
  - The closer r is to 0, the weaker the linear relationship between the variables.

DO THE MATH

PLANNING NOTES

Activity 2: Matching Correlation Coefficients (*Optional, 10 minutes*)

Instructional Routines: Take Turns; Collect and Display (MLR2); Discussion Supports (MLR 8) - Responsive Strategy	
Addressing: NC.M1.S-ID.8	

In this optional activity, students gain a better understanding of correlation coefficients by *Taking Turns* with a partner to match scatter plots and correlation coefficients. Students trade the roles of explaining their thinking and listening, providing opportunities to both explain their reasoning and critique the reasoning of others (MP3).

RESPONSIVE STRATEGIES

Display or provide charts with symbols and meanings. Create a chart of the r value ranging from -1 to 1. Label with the corresponding features of the various domains of the r value. Select students to label the chart with the corresponding descriptors for each domain. Small sketches or print outs of example scatter plots can be added to the appropriate areas of the chart.

Supports accessibility for: Conceptual processing; Memory

Step 1

- Keep students in pairs.
- Tell students that for each scatter plot, one partner finds the associated correlation coefficient and explains why they think it goes with that scatter plot. The other partner's job is to listen and make sure they agree. If they don't agree, the partners discuss until they come to an agreement. For the next scatter plot, the students swap roles. If necessary, demonstrate this protocol before students start working.

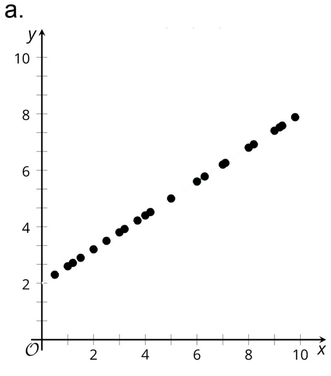
Advancing Student Thinking: Students may struggle with starting to match the scatter plots with a correlation coefficient. Guide students by asking them about the sign of the correlation coefficients. Ask them to sort the cards into groups that make sense and use that to connect to the correlation coefficient values. Ask them: "How does the sign of the correlation coefficient relate to the linear model?"

Student Task Statement

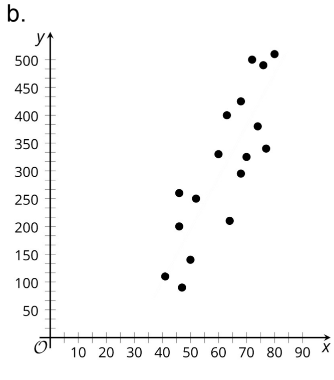
1. Take turns with your partner to match a scatter plot with a correlation coefficient.
2. For each match you find, explain to your partner how you know it's a match.
3. For each match your partner finds, listen carefully to their explanation. If you disagree, discuss your thinking and work to reach an agreement.

r = -1	r = -0.95	r = -0.74	r = -0.06	r = 0.48	r = 0.65	r = 0.9	r = 1

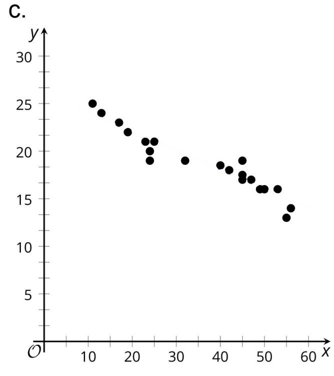
a.



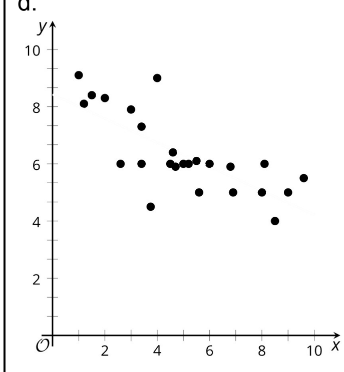
b.

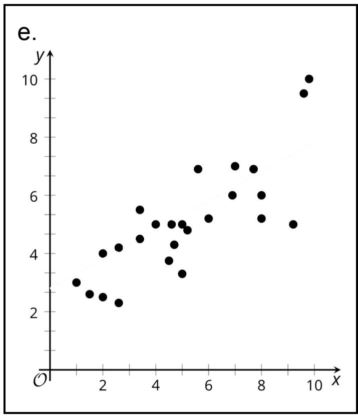


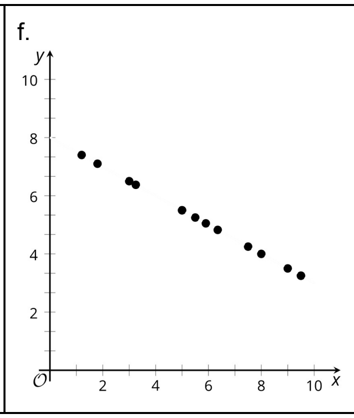
c.

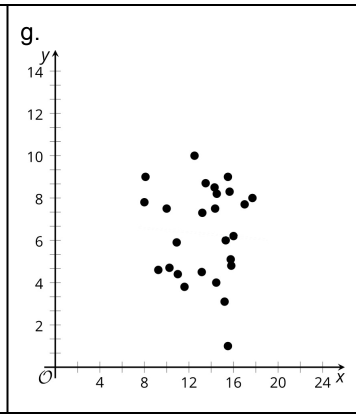


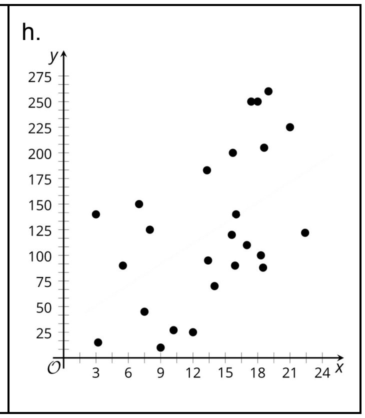
d.











Step 2

- Use the *Collect and Display* routine to facilitate a whole-class discussion. The purpose of this discussion is for students to understand that the correlation coefficient is a formal way to quantify the strength of a linear relationship between variables, and that the sign of the correlation coefficient tells you whether or not the variables show a positive or negative relationship. Below are some questions for discussion. Scribe student responses in a place for all to see, particularly with the second question. Listen for students' ideas about the meaning and implications of a correlation coefficient of 1 or -1.
  - "What does the sign of the correlation coefficient tell you about the data?" (If it is negative, then y tends to decrease as x increases. If it is positive, then y tends to increase as x increases.)
  - "What does it mean to have a correlation coefficient of 1 or -1?" (It means that the data set is perfectly linear, or one thing determines the other exactly, or two things are perfectly correlated, or the line is a perfect fit.)

RESPONSIVE STRATEGY

Use this routine to support whole-class discussion. Each time a student shares their ideas about the meaning of the correlation coefficient, press for details by requesting students to challenge an idea, elaborate on an idea, or give an example. Revoice student responses to demonstrate the relevance of mathematical language. Call students' attention to any words or phrases that helped clarify the original statement. If needed, practice phrases or words through choral response. This provides more students with an opportunity to produce language as they share their ideas about the relationship between data represented in scatter plots and correlation coefficients.

Discussion Supports (MLR8)

DO THE MATH

PLANNING NOTES

Activity 3: Copious Correlations (*10 minutes*)

Addressing: NC.M1.S-ID.6, NC.M1.S-ID.8

The mathematical purpose of this activity is to use the correlation coefficient to describe the relationship between two variables. Students examine a pair of variables and a correlation coefficient to describe the relationship between the variables as strong or weak and as positive or negative. Students must reason abstractly and quantitatively (MP2) when they interpret the situation to describe the relationship.

Step 1

- Keep students in pairs.
- Display the following example of the type of response expected. Have students reference it in their Student Workbooks to highlight or annotate to anchor understanding.

Student Task Statement

Example response:

> Cost of a package of light bulbs and number of lightbulbs in the package. $r = 0.96$
>
> Sample Response: The cost of a package of light bulbs and the number of light bulbs in the package has a correlation coefficient value near 1. This means that these variables have a very strong, positive linear relationship. In other words, the price of the package is very closely related to the number of light bulbs in the package (strong relationship) and when one of the variables goes up, the other variable tends to go up, too (positive relationship).

For each situation, describe the relationship between the variables, based on the correlation coefficient. Make sure to mention whether there is a strong relationship or weak relationship as well as whether it is a positive relationship or negative relationship.

1. Number of steps taken per day and number of kilometers walked per day. $r = 0.92$
2. Temperature of a rubber band and distance the rubber band can stretch. $r = 0.84$
3. Car weight and distance traveled using a full tank of gas. $r = -0.86$
4. Average fat intake per citizen of a country and average cancer rate of a country. $r = 0.73$
5. Score on science exam and number of words written on the essay question. $r = 0.28$
6. Average time spent listening to music per day and average time spent watching TV per day. $r = -0.17$

Step 2

- Facilitate a whole-class discussion. The purpose of this discussion is for students to interpret the data based on the relationship between the two variables that they determined using the correlation coefficient. Here are some questions for discussion:
  - "Do these correlations make sense based on what you understand about these variables?" (Yes, mostly. Walking more, for example, usually means you take more steps.)
  - "What does it mean for the relationship between the score on a science exam and the number of words written to be weak and positive?" (It means that, in general, the exam score tended to increase as the number of words written increased, but the relationship between these two variables was weak. This makes sense, because if you wrote very few words, you probably would not get a good score, but writing lots of words does not necessarily guarantee you a good score.)
  - "Which relationships would you consider modeling with a line? (1, 2, 3, and 4. The other two are too weak.)

DO THE MATH	PLANNING NOTES

Lesson Debrief (*5 minutes*)

The purpose of this lesson is for students to interpret the correlation coefficient and describe the relationship between two variables.

Choose what questions to focus the discussion on, whether students should first have an opportunity to reflect in their workbooks or talk through these with a partner, and what questions will be prioritized in the full class discussion. Consider providing students with whiteboards to share their sketches.

- "What might a scatter plot look like when the relationship between the two variables has a correlation coefficient of 0.9? Sketch it." (It looks like points that have a strong linear shape. It will have a positive direction.)
- "What does a scatter plot look like when the relationship has a correlation coefficient of -0.5? Sketch it." (It looks like a loosely scattered cloud of data that trends downward from left to right.)
- "What does it mean for two variables to have a weak, positive relationship?" (When one of the variables increases, the other variable also tends to increase, but since the linear relationship is weak, the data do not follow a linear path.)
- "What does the correlation coefficient tell you about the relationship between two variables?" (It tells you if the two variables are positively or negatively related. It also tells you how strong the linear relationship is between the two variables.)

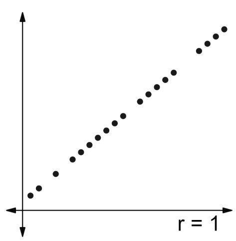
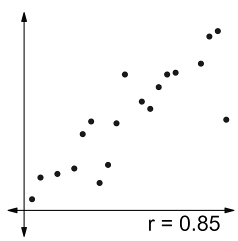
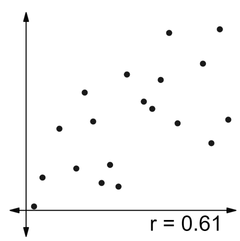
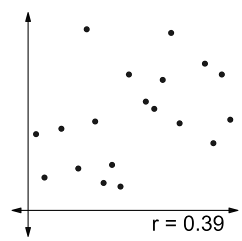
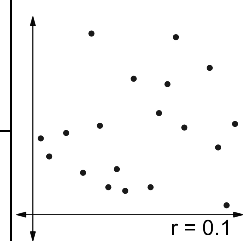
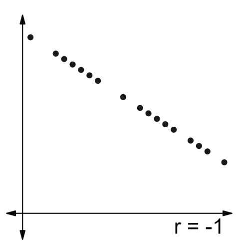
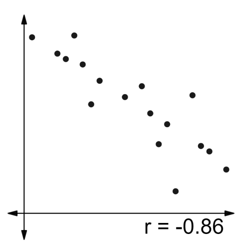
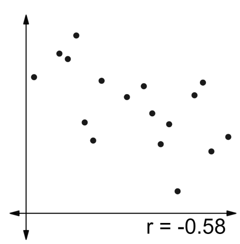
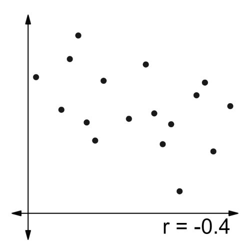
PLANNING NOTES

Student Lesson Summary and Glossary

While scatter plots can be used to identify the shape and direction of a set of data, we need a way to determine the strength of a linear relationship.

The **correlation coefficient** is a number that can be used to describe the strength and direction of a linear relationship. Usually represented by the letter r, the correlation coefficient can take values from -1 to 1. The sign of the correlation coefficient indicates the direction of the relationship. The closer the correlation coefficient is to 1 or -1, the stronger the linear relationship. The closer the correlation coefficient is to 0, the weaker the linear relationship.

> **Correlation coefficient:** A number between -1 and 1 that describes the strength and direction of a linear relationship between two numerical variables. The sign of the correlation coefficient indicates the direction of the relationship.The closer the correlation coefficient is to 1 or -1, the stronger the linear relationship. The closer the correlation coefficient is to 0, the weaker the linear relationship.

	Perfect correlation	Strong correlation	Moderate correlation	Weak correlation	No correlation
Positive					
Negative					

A correlation coefficient with a value near 1 suggests a strong, positive linear relationship between the variables. This means that most of the data tend to be tightly clustered into the shape of a line, and that when one of the variables increases in value, the other does as well. The number of schools in a community and the population of the community are examples of variables that have a strong, positive correlation. When there is a large population, there is usually a large number of schools, and small communities tend to have fewer schools, so the correlation is positive. These variables are closely tied together, so the correlation is strong.

Similarly, a correlation coefficient near -1 suggests a strong, negative relationship between the variables. Again, most of the data tend to be tightly clustered into the shape of a line, but now, when one value increases, the other decreases. The time since you left home and the distance left to reach school has a strong, negative correlation. As the travel time increases, the distance to school tends to decrease, so this is a negative correlation. The variables are again closely, linearly related, so this is a strong correlation.

Weaker correlations mean there may be other reasons the data are variable other than the connection between the two variables. For example, the number of pets and number of siblings has a weak correlation. There may be some relationship, but there are many other factors that account for the variability in the number of pets other than the number of siblings.

The context of the situation should be considered when determining whether the correlation value is strong or weak. In physics, measuring with precise instruments, a correlation coefficient of 0.8 may not be considered strong. In social sciences, collecting data through surveys, a correlation coefficient of 0.8 may be very strong.

Cool-down: How Bad Is It, Doc? (*5 minutes*)

Addressing: NC.M1.S-ID.8

Cool-down Guidance: More Chances
In the next couple of lessons, students will use technology to calculate the correlation coefficient and interpret it to determine if a linear model is appropriate to use for a set of data.

Cool-down

Doctors suspect a strain of bacteria found in the hospital is becoming resistant to antibiotics. They put various concentrations of an antibiotic in petri dishes and add some of the bacteria to allow it to grow. The bacteria grow into groups in the dish called colonies. After some time, the doctors return to the petri dishes and count the number of colonies for the different amounts of antibiotic.

The correlation coefficient of the data is $r = -0.83$.

1. What does the sign of the correlation coefficient tell you about the relationship between the number of bacteria colonies and the concentration of antibiotic in the dish?
2. What does the numerical value of the correlation coefficient tell you about the relationship between the number of bacteria colonies and the concentration of antibiotic in the dish?
3. In a follow-up study, a group of scientists collect data with a correlation coefficient of $r = -0.94$. Which study suggests a stronger relationship between the number of bacteria colonies and the concentration of the antibiotic—the doctors' study or the scientists' study? Explain your reasoning.

Student Reflection:

When we work in groups in math class I feel:

a. Excited b. It does not matter to me c. Not excited

I feel this way because ___________________________.

DO THE MATH

INDIVIDUAL STUDENT DATA

SUMMARY DATA

NEXT STEPS

TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

With which math ideas from today's lesson did students grapple most? Did this surprise you or was this what you expected?

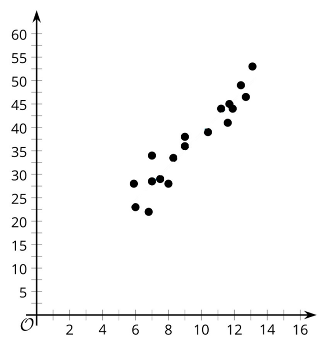
Practice Problems

1. The correlation coefficient, r, is given for several different data sets. Which value for r indicates the **strongest** linear relationship?

   a. 0.01
   b. -0.34
   c. -0.82
   d. -0.95

2. Which of the values is the **best** estimate of the correlation coefficient for the data shown in the scatter plot?

   a. -0.9
   b. -0.4
   c. 0.4
   d. 0.9



3. The number of hours worked, x, and the total dollars earned, y, have a strong positive relationship.

   Explain what it means to have a strong positive relationship in this situation.

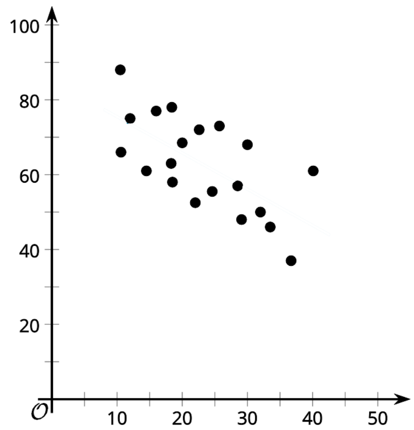
4. The number of minutes on the phone and the customer satisfaction rating have a weak negative relationship.

   Explain what it means to have a weak negative relationship in this context.

5. The correlation coefficient, r, is given for several different data sets. Which value for r indicates the **weakest** linear relationship?

   a. 0.01
   b. 0.5
   c. -0.99
   d. 1

6. Which of the following is the **best** estimate of the correlation coefficient for the relationship shown in the scatter plot?



a. -0.9
b. -0.4
c. 0.4
d. 0.9

7. (*Technology required.*)

a. Use the data in the table to make a scatter plot.[1] Data can be accessed through this spreadsheet: https://bit.ly/U4L2DataSet.

b. Describe the relationship between the variables.

c. What does the point $(62, 1320)$ represent?

(From Unit 4, Lesson 1)

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	1320
Rhesus monkey	7	179
Kangaroo	35	56
Sheep	56	175
Jaguar	100	157
Chimpanzee	52	440
Pig	192	180

8. For a baseball team fundraiser, Noah washed 16 cars and earned $213 for the team! He charged $18 per car wash.

a. Write a linear equation in point-slope form representing the amount of money, y, a player could raise for washing x cars.

b. Write a linear equation in slope-intercept form representing the amount of money, y, a player could raise for washing x cars.

c. What could the y-intercept of the equation represent?

(From Unit 3)

9. Solve the equation: $\frac{3}{5}x - 6 = \frac{1}{3}x + 4$

(From Unit 2)

[1] Adapted from IM 6–8 Math https://curriculum.illustrativemathematics.org/MS/index.html, which was originally developed by Open Up Resources and authored by Illustrative Mathematics, and is copyright 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). OUR's 6–8 Math Curriculum is available at https://openupresources.org/math-curriculum/. Adaptations and updates to IM 6–8 Math are copyright 2019 by Illustrative Mathematics, and are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Lesson 3: Linear Models

PREPARATION

Lesson Goals	Learning Targets
• Informally assess the goodness of fit for a given linear model to a scatter plot of data. • Interpret (orally and in writing) the rate of change and vertical intercept for a linear model in everyday language. • Predict (extrapolate) and estimate (interpolate) values not given in the data set by using an equation representing the linear model.	• I can assess if a linear model fits the data well and use the linear model to estimate values I want to find. • I can describe the rate of change and vertical intercept (y-intercept) for a linear model in everyday language.

Lesson Narrative

In this lesson, students will work to understand how a linear model is used to describe the relationship between two numerical variables, to recognize and describe characteristics of a best fit line, and to use a line of best fit to make predictions. The work of this lesson connects to previous work because students investigated patterns of association in bivariate data in eighth grade and in previous lessons. The work of this lesson connects to upcoming work because students will use technology to find the line of best fit for the data and use it to solve problems.

When students articulate things they notice and things they wonder about a scatter plot and the accompanying linear model, they have an opportunity to attend to precision in the language they use to describe what they see (MP6). Students reason abstractly and quantitatively by making sense of slope and intercept in context (MP2).

Where do you see connections between material students shared and discussed in previous lessons and the material in this lesson?

Focus and Coherence

Building On	Addressing	Building Towards
NC.8.SP.2: Model the relationship between bivariate quantitative data to: ● Informally fit a straight line for a scatter plot that suggests a linear association. ● Informally assess the model fit by judging the closeness of the data points to the line. **NC.8.SP.3** Use the equation of a linear model to solve problems in the context of bivariate quantitative data, interpreting the slope and y-intercept.	**NC.M1.S-ID.6:** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related. **NC.M1.S-ID.7:** Interpret in context the rate of change and the intercept of a linear model. Use the linear model to interpolate and extrapolate predicted values. Assess the validity of a predicted value.	**NC.M1.S-ID.6a:** Fit a least squares regression line to linear data using technology. Use the fitted function to solve problems.

Agenda, Materials, and Preparation

- **Bridge** (*Optional, 5 minutes*)
- **Warm-up** (*5 minutes*)
- **Activity 1** (*15 minutes*)
  - Technology is required for this activity: Acquire devices that can access Desmos (recommended) or other graphing technology. It is ideal if each student has their own device.
  - Vimeo video “Oranges in a Box:” https://bit.ly/OrangesInABox
- **Activity 2** (*10 minutes*)
- **Lesson Debrief** (*5 minutes*)
- **Cool-down** (*5 minutes*)
  - M1.U4.L3 Cool-down (print 1 copy per student)

LESSON

Bridge (*5 minutes*)

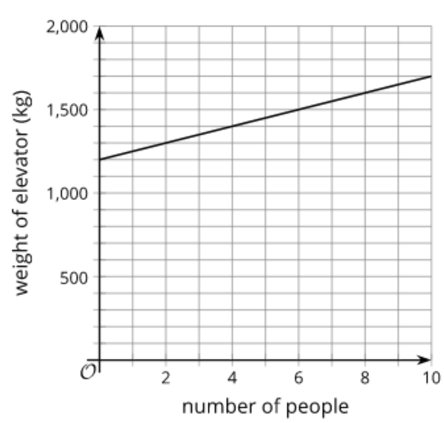
Building From: NC.8.SP.3

This bridge helps students understand that the slope and y-intercept of a linear model can be used to interpret a situation. This task is aligned to question 4 in Check Your Readiness.

Student Task Statement

Here is a linear model of the weight of an elevator and the number of people on the elevator.

1. Find these values. Explain your reasoning.
   a. the weight of the elevator when 6 people are on it
   b. the number of people on the elevator when it weighs 1,400 kg
   c. the weight of the elevator when no people are on it
   d. the increase in elevator weight for each additional person according to the model
2. Which of your answers corresponds to the slope of the line in the graph?
3. Which of your answers corresponds to the *y*-intercept of the line in the graph?



Warm-up: Crowd Noise (*5 minutes*)

Instructional Routine: Notice and Wonder		
Building On: NC.8.SP.2	**Addressing:** NC.M1.S-ID.6	**Building Towards:** NC.M1.S-ID.7

The purpose of this warm-up is to help students recall information about fitting lines to data in scatter plots, which will be useful when students expand their understanding in a later activity.

While students may notice and wonder many things about these images, the relationship between the number of people and the maximum noise level, the interpretation of the line of best fit, and a general understanding of a scatter plot are the important discussion points.

Through articulating things they notice and things they wonder about the scatter plot and the linear model, students have an opportunity to attend to precision in the language they use to describe what they see (MP6). They might first propose less formal or imprecise language, and then restate their observation with more precise language in order to communicate more clearly.

Notice students who use correct terminology in their responses. In particular, the terms "linear model," "slope," and "vertical or y-intercept" are important to review during this warm-up.

Step 1

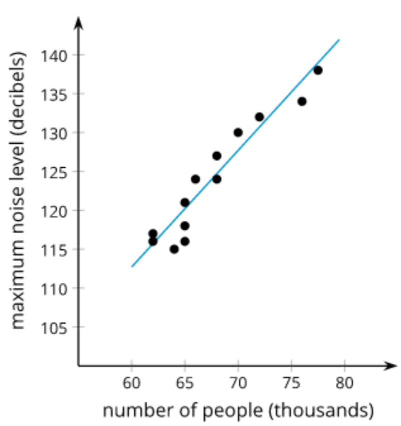
- Ask students to arrange themselves into pairs or use visibly random grouping.
- Display the image for all to see.
- Ask students what they *Notice and Wonder* about the image.
- Give students 1 minute of quiet think time and then 1 minute to discuss the things they notice and wonder with their partner, followed by a whole-class discussion.

Advancing Student Thinking: The vertical intercept appears to be approximately 105 decibels, but in fact it is a much smaller number: the origin is not shown on the graph.

Student Task Statement

What do you notice? What do you wonder?

$y = 1.5x + 22.7$



Step 2

- 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 on or near the image.
- After all responses have been recorded without commentary or editing, ask students: “Is there anything on this list you are wondering about now?” Encourage students to respectfully disagree, ask for clarification, or point out contradicting information.
- If students do not use these terms in their responses, ensure they recall the vocabulary of “slope” and “vertical or *y*-intercept.” Explain that a mathematical model is a way of representing data that may oversimplify but helps to make predictions and see patterns. Students made a model when planning a pizza party in Unit 2, Lesson 1. The line on the scatter plot is a linear model for the relationship between number of people and crowd noise.

DO THE MATH	PLANNING NOTES

Activity 1: Orange You Glad We’re Boxing Fruit? (*15 minutes*)

Instructional Routine: Co-Craft Questions (MLR5)		
Building On: NC.M1.S-ID.6	**Addressing:** NC.M1.S-ID.7	**Building Towards:** NC.M1.S-ID.6a

In this activity, students will use technology to create a scatter plot from data collected through a video simulation, informally assess a given linear model for how well it fits the data, interpret the slope and *y*-intercept of the linear model, and use the linear model to make predictions.

This activity works best when each student has access to devices that can run Desmos or other graphing technology because the level of precision is important. If students don’t have individual access, projecting the Desmos graph is helpful during the synthesis.

Step 1

- Ask students to arrange themselves into groups of two to four or use visibly random grouping.
- Play the video of adding oranges to a box on a scale. You may need to pause the video for students to write down the weights.
  - “Oranges in a Box” video available here: https://bit.ly/OrangesInABox.
- Using the *Co-Craft Questions* routine, ask students: “What mathematical questions could we ask about this situation?” Give each group 1 minute to generate a list of as many questions as they can about the situation. Then invite each group to share one question orally with the class that has not already been suggested. (Students do not need to answer any of these questions; the purpose is to provide an opportunity to produce the language of mathematical questions.)
- Have students work together to complete questions 1–3.

CO-CRAFT QUESTIONS	***What Is This Routine?*** Students are presented with a picture, video, diagram, data display, or description of a situation, and their job is to generate one or more mathematical questions that could be asked about the situation. Students then share and compare their questions, as the teacher calls attention to questions that align with the content goals of the lesson. Finally, the "official" question or problem is revealed for students to work on. ***Why This Routine?*** *Co-Craft Questions* (MLR5) allows students to get inside of a context before feeling pressure to produce answers, and creates space for students to produce the language of mathematical questions themselves. Use this routine to spark curiosity about a new mathematical idea or representation, and to elicit everyday student language to brainstorm about the quantitative relationships that might be investigated. During this routine, students use conversation skills and develop meta-awareness of the language used in mathematical questions and problems.

Advancing Student Thinking: Students may struggle with estimating a slope when the scales on the x and y axes are different. Ask students to find the coordinates for a couple of points on or near the line and find the slope between those points.

Student Task Statement

1. Watch the video and record the weight for the number of oranges in the box.

Number of oranges	3	4	5	6	7	8	9	10
Weight in kilograms								

2. Use technology to create a scatter plot of the data.
   a. Describe the relationship between the number of oranges and weight in kilograms.
   b. Which of the following is most likely the correlation coefficient of the data? Explain your reasoning.

      -0.92 -0.42 0.1 0.9

3. The linear equation $y = 0.216x + 0.345$ is suggested as a linear model for the data. Use technology to graph the line along with the scatter plot.
   a. How well does the line model the data? Explain your reasoning.
   b. What is the value of the slope, and what does it mean in this situation?
   c. What is the value of the y-intercept and what does it mean in this situation?
4. Use the line to predict the weight of a box containing 11 oranges. Will this estimate be close to the actual value? Explain your reasoning.
5. Use the line to predict the weight of a box containing 50 oranges. Will this estimate be close to the actual value? Explain your reasoning.

Step 2

- Invite students to share their interpretation of the slope and y-intercept. (Slope: for every additional orange placed in the box, the line predicts (or estimates) that the weight of the box will increase by 0.216 kg; on average, an orange weighs 0.216 kg. y-intercept: it's the weight of an empty box; it's the weight of a box with zero oranges.)
- If no one responds along the lines of "for each additional orange…" bring up this interpretation since it is the one that generalizes to other situations. Also emphasize that the slope has to do with the values predicted (or estimated) by the best fit line rather than the points in the table.
- Tell students to continue working with their group to complete questions 4 and 5.

Monitoring Tip: Monitor for strategies students use to make estimates.

- Use the graph of the linear model to find the points $(11, 2.72)$ and $(50, 11.145)$.
- Use the equation of the linear model to substitute the values of 11 and 50 for x and calculate the y value.

Step 3

- Tell students to keep this set of data for a future lesson.
- Facilitate a whole-class discussion. The purpose of the discussion is to share strategies for using the model to make predictions and to assess the validity of the predicted values.
- Select previously identified students to share their strategies from the Monitoring Tip. If one of the strategies is not mentioned, bring it up. Here are some questions to discuss these strategies:
  - "Are the values obtained from looking at the scatter plot close to the values calculated using the equation?" (Yes, the values are close.)
  - "Which method will you use to predict values based on the model?" (Maybe a mixture of both methods. Using the equation produces a more precise value from the model, but it is also important to look at the graph to get a sense of how well the model fits the data where I am predicting.)
- Here are some questions to discuss making predictions:
  - "Why are the estimates considered predictions and not actual values?" (There weren't 11 oranges or 50 oranges actually put into the box, so the estimates are only predictions for what the box might weigh.)
  - "What's the difference between the predicted weight of the box and the actual weight when there were 5 oranges?" (The linear model predicts 1.425 kilograms with 5 oranges, but the weight was actually 1.502 kilograms. This means the prediction was under by 0.077 kilograms.)
- Here are some questions to discuss validity:
  - "Which prediction do you feel is more accurate for the data collected: the prediction for 11 oranges or for 50 oranges? Explain your reasoning." (I feel more confident in the prediction for 11 oranges. I think one more orange will continue the linear pattern. I feel less confident in the 50 oranges. There are too many unknowns. Will 50 oranges even fit in the box? Will the sizes of the oranges vary more, smaller or larger?)
  - "How would the scatter plot and linear model change if grapefruits were used instead of oranges?" (The weight would be increased for each point, and the slope of the line would be greater.)
  - "How would the scatter plot and linear model change if the box itself was heavier?" (The points and line would shift up.)

DO THE MATH	PLANNING NOTES

Activity 2: The Slope Is the Thing (*10 minutes*)

Instructional Routine: Discussion Supports (MLR8) - Responsive Strategy	
Building On: NC.M1.S-ID.6	**Addressing:** NC.M1.S-ID.7

The purpose of this activity is for students to interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data. Students are given scatter plots for different pairs of variables and the equation of a line of best fit. Students use the line of best fit and its equation to describe the meaning of the vertical intercept and slope.

Step 1

- Ask students to arrange themselves into pairs or use visibly random grouping.
- Provide students a few minutes of quiet think time for questions 1 and 2 before comparing responses with their partner. If time is limited, ask partners to choose one positive relationship and one negative relationship they will both describe.
- Ask students to compare their responses to their partner's for the scatter plots and decide if both their responses are correct for each scatter plot, even if they are different.

RESPONSIVE STRATEGIES

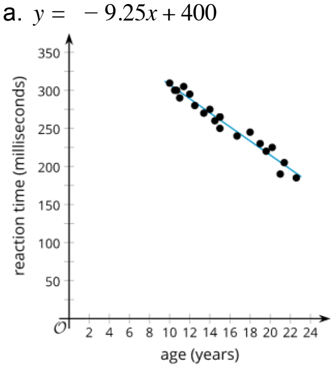
To help students get started, provide sentence frames such as: "What does this part of ___ mean?" and "I predict ___ because . . . ." Encourage students to annotate their graphs while using the sentence frames. Ask students to take turns sharing their interpretation of the slope and vertical intercept for each scatter plot. Display the following sentence frames for all to see: "___ represents ___", "Is there another way to say . . .?", and "It looks like ___ represents . . . ." Encourage students to challenge each other when they disagree. This will help students clarify their understanding of the meaning of slope and the vertical intercept in context.

Discussion Supports (MLR8)

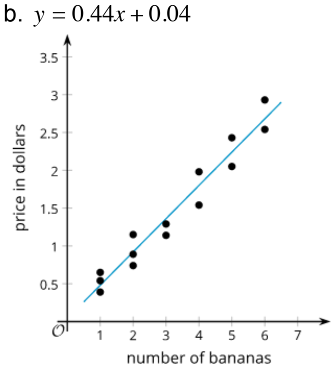
Student Task Statement

Here are several scatter plots. The linear models given here are lines of best fit.

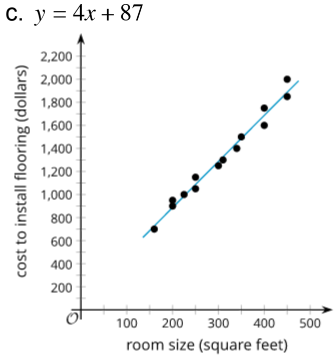
a. $y = -9.25x + 400$



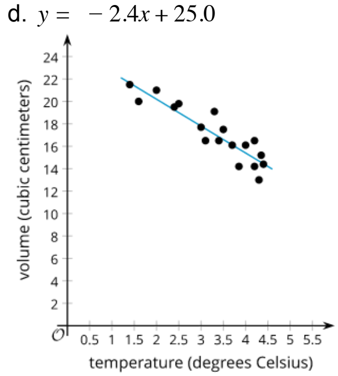
b. $y = 0.44x + 0.04$



c. $y = 4x + 87$



d. $y = -2.4x + 25.0$

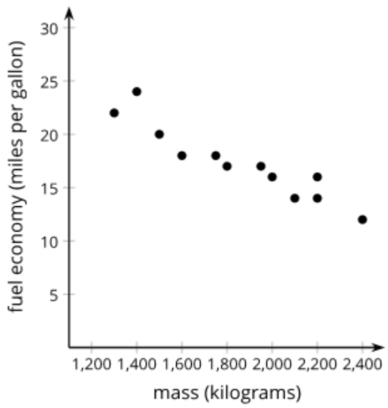


1. Using the horizontal axis for x and the vertical axis for y, interpret the slope of each linear model in the situations shown in the scatter plots.
2. Identify and interpret the y-intercept of each linear model in the situations provided.

Are You Ready For More?

Clare, Diego, and Elena collect data on the mass and fuel economy of cars at different dealerships. Clare finds the line of best fit for data she collected for 12 used cars at a used car dealership. The line of best fit is $y = \frac{-9}{1000}x + 34.3$, where x is the car's mass, in kilograms, and y is the fuel economy, in miles per gallon.

Diego made a scatter plot for the data he collected for 10 new cars at a different dealership.



Elena made a table for data she collected on 11 hybrid cars at another dealership.

Mass (kilograms)	Fuel economy (miles per gallon)
1,100	38
1,200	39
1,250	35
1,300	36
1,400	31
1,600	27
1,650	28
1,700	26
1,800	28
2,000	24
2,050	22

1. Interpret the slope and y-intercept of Clare's line of best fit in this situation.
2. Diego looks at the data for new cars and used cars. He claims that the fuel economy of new cars decreases as the mass increases. He also claims that the fuel economy of used cars increases as the mass increases. Do you agree with Diego's claims? Explain your reasoning.
3. Elena looks at the data for hybrid cars and correctly claims that the fuel economy decreases as the mass increases. How could Elena compare this decrease for hybrid cars to the decrease for new cars? Explain your reasoning.

Step 2

- Facilitate a whole-class discussion. The purpose of this discussion is for students to describe the rate of change and the vertical intercept using the context in each graph.
- For each question, give students time to think individually and then share their response with their partner. Select a student or pair of students to respond to the question. Ask students:
  - "Why is the intercept for the bananas not $(0,0)$?" (A linear model is not exact even for the data it is based on. It is an approximation based on the data in the scatter plot. It is possible that it represents the weight of the bag that bananas were placed in. It is also possible that this value does not make sense since there is no evidence to believe the same linear relationship will hold near zero.)
  - "How do you interpret the slope for each equation?" (The slope is the change in y divided by the change in x, so look at the labels on the scatter plot and talk about how much the *y*-variable increases or decreases when the *x*-variable increases by one. If the slope is negative, then on average the *y*-variable will decrease with an increase in *x*. If the slope is positive, then on average the *y*-variable will increase when *x* does.)

- "When might it make sense to interpret the y-intercept for a linear model?" (When the data used to create the model has *x*-values near zero. In other cases, we should not put too much faith in the answer since the linear trend may not continue farther from the collected data.)

DO THE MATH	PLANNING NOTES

Lesson Debrief (*5 minutes*)

The goal of this discussion is for students to make connections between bivariate data, a linear model, and the context of the data.

Choose what questions to focus the discussion on, whether students should first have an opportunity to reflect in their workbooks or talk through these with a partner, and what questions will be prioritized in the full class discussion. Consider providing students with whiteboards to share their sketches.

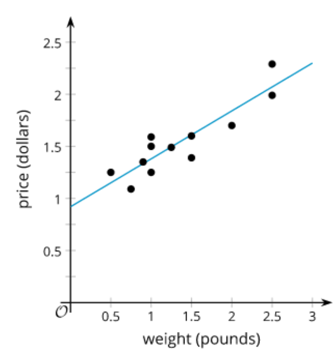
- "Why is creating a linear model useful?" (A linear model allows us to make predictions for the data in a range near the given data. It also helps to describe the relationship between the two variables quantitatively.)
- "What are some situations in which you have encountered a scatter plot or a line of best fit previously? What is the meaning of the slope and vertical intercept of the line of best fit in this context?" (In science class, we graphed the relationship between the temperature and the time it took for a reaction to take place. The slope represents how much the reaction time changes, on average, for each unit increase in temperature. The vertical intercept represents the reaction time when the temperature is 0.)

PLANNING NOTES

Student Lesson Summary and Glossary

While working in math class, it can be easy to forget that reality is somewhat messy. Not all oranges weigh exactly the same amount, beans have different lengths, and even the same person running a race multiple times will probably have different finishing times. We can approximate these messy situations with more precise mathematical tools to better understand what is happening. We can also predict or estimate additional results as long as we continue to keep in mind that reality will vary a little bit from what our mathematical model predicts.

For example, the data in this scatter plot represent the price of a package of broccoli and its weight. The data can be modeled by a line given by the equation $y = 0.46x + 0.92$. The data points do not all fall on the line because there may be factors other than weight that go into the price, such as the quality of the broccoli, the region where the package is sold, and any discounts happening in the store.



We can interpret the y-intercept of the line as the price for the package without any broccoli (which might include the cost of things like preparing the package and shipping costs for getting the vegetable to the store). In many situations, the behavior may not follow the same linear model farther away from the given data, especially as one variable gets close to zero. For this reason, the interpretation of the y-intercept should always be considered in context to determine if it is reasonable to make sense of the value in that way.

We can also interpret the slope as the approximate increase in price of the package for the addition of 1 pound of broccoli to the package.

The equation also allows us to predict additional values for the price of a package of broccoli for packages that have weights near the weights observed in the data set. For example, even though the data does not include the price of a package that contains 1.7 pounds of broccoli, we can predict the price to be about $1.70 based on the equation of the line, since $0.46 \cdot 1.7 + 0.92 \approx 1.70$.

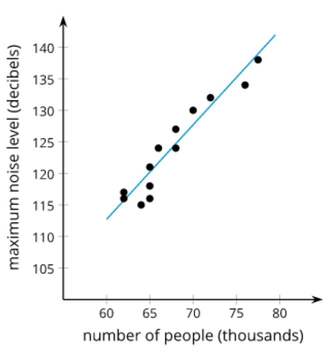
On the other hand, it does not make sense to predict the price of 1,000 pounds of broccoli with these data points, because there may be many more factors that will influence the pricing of packages with weights so much larger than the data points presented here.

Cool-down: Roar of the Crowd (*5 minutes*)

Building On: NC.M1.S-ID.6	**Addressing:** NC.M1.S-ID.7
Cool-down Guidance: Points to Emphasize Question 3 provides a great opportunity to highlight that we have to be careful when estimating the y-intercept visually, since the vertical axis is not always drawn where $x = 0$. In these cases, it is often better to find the y-intercept using the regression equation.	

Cool-down

Here is the scatter plot from today's warm-up.The scatter plot shows the maximum noise level when different numbers of people are in a stadium. The linear model is given by the equation $y = 1.5x + 22.7$, where y represents maximum noise level and x represents the number of people, in thousands, in the stadium.



1. The slope of the linear model is 1.5. What does this mean in terms of the maximum noise level and the number of people?

2. A sports announcer states that there are 65,000 fans in the stadium. Estimate the maximum noise level. Is this estimate reasonable? Explain your reasoning.

3. What is the y-intercept of the linear model given? What does it mean in the context of the problem? Is this reasonable? Explain your reasoning.

Student Reflection:

Today's lesson called for a great deal of sharing your thinking and strategies. On a scale of 1 to 5, how confident were you to speak aloud and share with your peers?

1 - Not confident at all
2 - Barely confident
3 - Not completely confident but not anxious
4 - Fairly confident
5 - Very confident

How can your teacher support you in becoming more confident?

DO THE MATH

INDIVIDUAL STUDENT DATA

SUMMARY DATA

NEXT STEPS

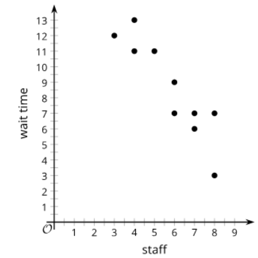
TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

How effective were your questions in supporting students' thinking today? What did students say or do that showed they were effective?

Practice Problems

1. The scatter plot shows the number of minutes people had to wait for service at a restaurant and the number of staff available at the time.

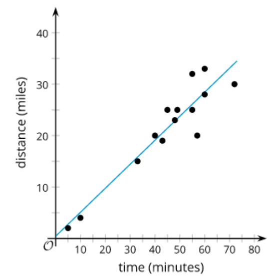


   A line that models the data is given by the equation $y = -1.62x + 18$, where y represents the wait time, and x represents the number of staff available.

   a. The slope of the line is -1.62. What does this mean in this situation? Is it realistic?

   b. The y-intercept is $(0, 18)$. What does this mean in this situation? Is it realistic?

2. A taxi driver records the time required to complete various trips and the distance for each trip.



   The best fit line is given by the equation $y = 0.467x + 0.417$, where y represents the distance in miles, and x represents the time for the trip in minutes.

   a. Use the best fit line to predict the distance for a trip that takes 20 minutes. Show your reasoning.

   b. Use the best fit line to predict the time for a trip that is 6 miles long. Show your reasoning.

3. (*Technology required.*) The table below represents the height of students, in inches, in Ms. Maas' Math 2 class with their corresponding scores on the last test.

Height (inches)	58	61	61	63	64	67	68	68	72	72	73	74
Test score (out of 100)	80	94	72	75	88	96	89	90	73	84	100	87

   a. Use technology to make a scatter plot.

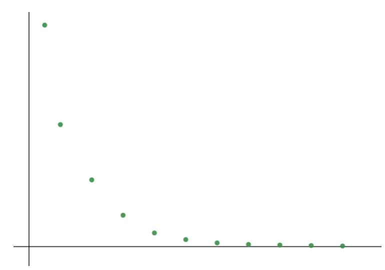
   b. Is there a strong or weak relationship between the students' heights and test scores?

   c. Is there a positive or negative relationship between the students' heights and test scores?

   d. Did you expect a different answer? Why or why not?

   (From Unit 4, Lesson 2)

4. Does the scatter plot represent:



   a. A linear or nonlinear relationship?
   b. A positive or negative relationship? Or neither?

   (From Unit 4, Lesson 1)

5. Parallelogram $ABCD$ has vertices at $A(-3, 6.5)$, $B(-1, -1.5)$, and $C(4.5, 0)$. What are the coordinates of Point D?

   (From Unit 3)

6. In a video game, players form teams and work together to earn as many points as possible for their team. Each team can have between two and four players. Each player can score up to 20 points in each round of the game. Han and three of his friends decided to form a team and play a round.

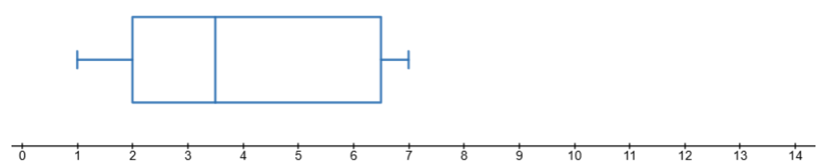
   Write an expression, an equation, or an inequality for each quantity described here. If you use a variable, specify what it represents.

   a. the allowable number of players on a team
   b. the number of points Han's team earns in one round if every player earns a perfect score
   c. the number of points Han's team earns in one round if no players earn a perfect score
   d. the number of players in a game with six teams of different sizes: two teams have four players each, and the rest have three players each
   e. the possible number of players in a game with eight teams

   (From Unit 2)

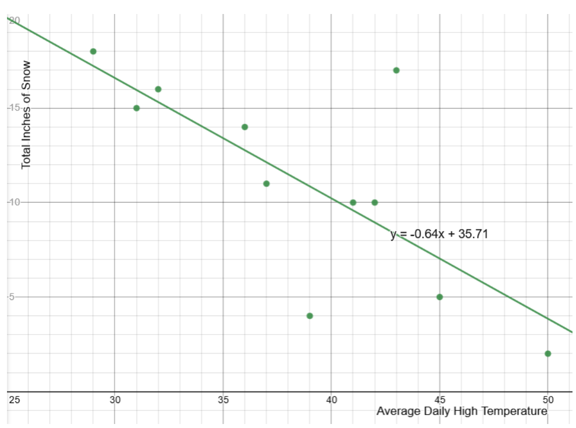
7. The box plot below represents a data set with a maximum value of 7. If the maximum value were adjusted:

   a. What is the smallest whole number it could be so that the whisker to the right would be longer than the box to the right?
   b. How would adjusting the maximum value affect the median? How do you know?



   (From Unit 1)

8. The scatter plot and line of best fit below represent the average daily high temperature in winter (x) and total snowfall in inches (y) for several cities one winter. In Charlotte, the average high temperature one winter was 51^0 F. Based on the linear relationship, how much snow would you expect Charlotte to get that winter?



   (Addressing NC.8.SP.3)

Lesson 4: Fitting Lines

PREPARATION

Lesson Goal	Learning Targets
• Use technology to generate the line of best fit and calculate the correlation coefficient.	• I can use technology to find the line of best fit. • I can use technology to calculate the correlation coefficient.

Lesson Narrative

This lesson builds on students' understanding of best fit lines developed in the previous lesson. Here, students use technology to find equations for lines of best fit, known as regression equations. They can now take any data set, find the regression equation, and use that equation to make predictions for x-values not in the data set. They can interpret the slope and vertical intercept in context. Additionally, students return to the idea, raised briefly in the previous lesson, that even models that fit the data well do not make good predictions everywhere. They learn that **interpolation**—predictions for values within the boundaries of the data set—is generally reliable, while **extrapolation**—predictions for values outside these boundaries—may not be. Students will access data from a spreadsheet that allows them to copy and paste as needed; data points are provided on different tabs of the spreadsheet aligned with Activity 1, Activity 2, Are You Ready For More?, and the Cool-down: https://bit.ly/U4L4DataSet.

Students are reasoning abstractly and quantitatively (MP2) when they interpret the meaning of the slope and vertical intercept in context. When they assess whether a value predicted by a model is reasonable, they are constructing viable arguments and critiquing the reasoning of others (MP3).

What is the main purpose of this lesson? What is the one thing you want your students to take away from this lesson?

Focus and Coherence

Building On	Addressing
NC.8.SP.2: Model the relationship between bivariate quantitative data to: ● Informally fit a straight line for a scatter plot that suggests a linear association. ● Informally assess the model fit by judging the closeness of the data points to the line.	**NC.M1.S-ID.6:** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related. **a**. Fit a least squares regression line to linear data using technology. Use the fitted function to solve problems. **NC.M1.S-ID.7:** Interpret in context the rate of change and the intercept of a linear model. Use the linear model to interpolate and extrapolate predicted values. Assess the validity of a predicted value.

Agenda, Materials, and Preparation

A data spreadsheet will be used in Activity 1, Activity 2, Are You Ready For More?, and the Cool-down of this lesson. Students can access this data through this link: https://bit.ly/U4L4DataSet.

- **Warm-up** (*5 minutes*)
- **Activity 1** (*15 minutes*)
  - Technology is required for this lesson: Acquire devices that can access Desmos (recommended) or other graphing technology. It is ideal if each student has their own device.
- **Activity 2** (*15 minutes*)
- **Lesson Debrief** (*5 minutes*)
- **Cool-down** (*5 minutes*)
  - M1.U4.L3 Cool-down (print 1 copy per student)

LESSON

Warm-up: Selecting the Best Line (*5 minutes*)

Building On: NC.8.SP.2	**Building Towards:** NC.M1.S-ID.6a

The mathematical purpose of this activity is for students to be able to visually assess the best line that fits data points among a set of choices. Students are given a scatter plot and two lines that may fit the data. Students must select the line that better fits the data. The given lines address these common errors in students thinking about best fit lines: going through the most points, dividing the data in half, and connecting the points on both ends of the scatter plot.

Listen for students using the terms “slope” and “y-intercept.”

Step 1

- Ask students to arrange themselves in pairs or use visibly random grouping. (Note: Students will remain in these partners for the duration of the lesson.)

- Display the graphs for all to see.

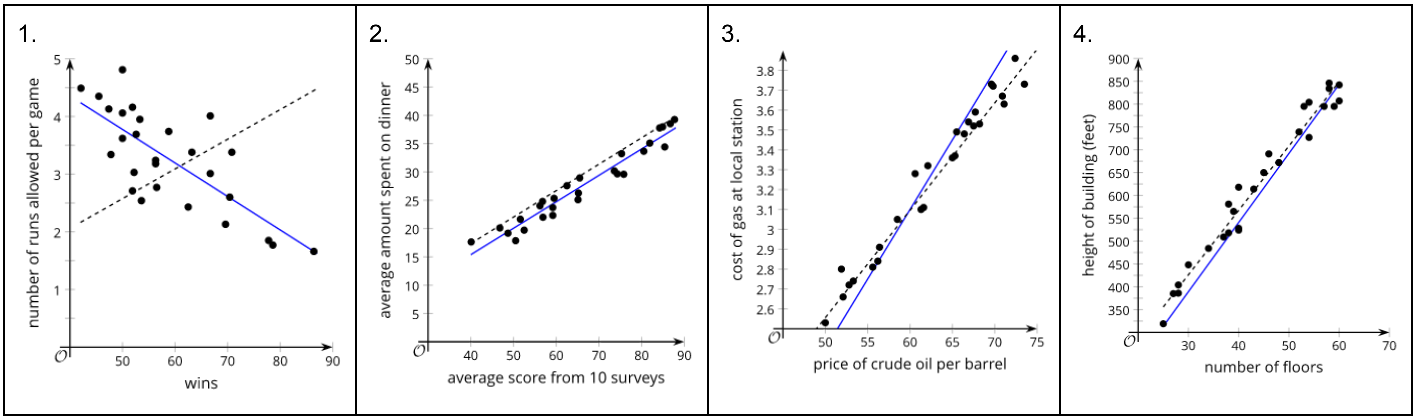
- Explain to students that each graph has two lines: a dashed line and a solid line. They must decide which line best fits the data.

- Give students 1 minute of quiet time followed by 2 minutes to share their thinking with their partner. After they have shared, move to the whole-class discussion.

> M **Monitoring Tip:** As students are sharing with their partner, listen for responses and justifications of the best fit line. Make note of any student groups that discuss slope or concepts of slope in their justification.

Student Task Statement

Which of the lines is the best fit for the data in each scatter plot? Explain your reasoning.



Step 2

- Ask the following questions to the class and consider having pre-identified partners (from Monitoring Tip) share.
  - “Is the dashed line in the question about runs and wins (#1) a bad fit, good fit, or best fit?” (The line is a bad fit because it does not show the correct relationship between the variables. It shows that the value of y increases as the value of x increases, rather than the value of y decreasing as the value of x increases.)
  - “Can you explain the relationship between the two lines in the questions about runs and wins using the concept of slope?” (The slope of the dashed line is positive, and the slope of the solid line is negative.)
  - “Is the dashed line in the question about oil and gas (#3) a bad fit, good fit, or best fit?” (It is the best fit because it is close to going through the middle of the data and follows the same trend as the data.)
  - “What factors helped you select the linear model that fits the data best?” (The line should go through the middle of the data, follow the trend of the data, and have a similar number of points on each side of the line.)
- According to the most often agreed-upon standard for measuring “fit” for data, the best fit line is the solid line in #1 and #2 and the dashed line in #3 and #4. What do you think this standard considers? (It goes through the center of the data. It gets close to all points.)

DO THE MATH

PLANNING NOTES

Activity 1: Fitting Lines with Technology (*15 minutes*)

Instructional Routines: Fit It; Critique, Correct, Clarify (MLR3)
Addressing: NC.M1.S-ID.6a

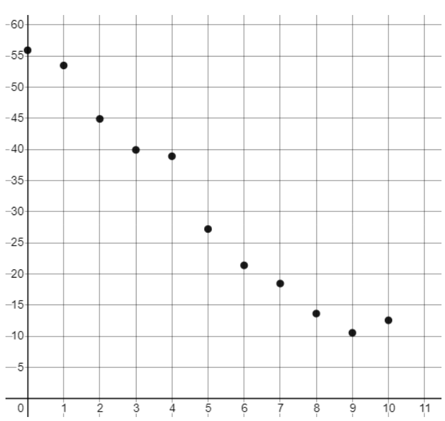
The mathematical purpose of this *Fit It* activity is for students to use technology to compute the equation for the line of best fit. This statistical process is called regression, and the resulting equation is referred to as a regression equation. The study of this process is not included in this course. Students in Math 1 should be familiar with the term "regression equation" and understand that it is the equation for the line of best fit. In addition to computing the regression equation, students use technology to find the correlation coefficient.

This is the first time in this course that students will participate in a *Fit It* routine.

FIT IT	***What Is This Routine?*** *Fit It* indicates activities where students have an opportunity to use a table of points to produce a graph to see patterns and make predictions. Also, when appropriate, students find a function that best fits the data. ***Why This Routine?*** Plotting values from a table into the coordinate plane provides opportunities to use mathematics to make relationships, trends, and patterns visible. Reasoning about functions that model data represented in these two ways (table and coordinate plane) supports students to recognize that functions can be built to fit real data.

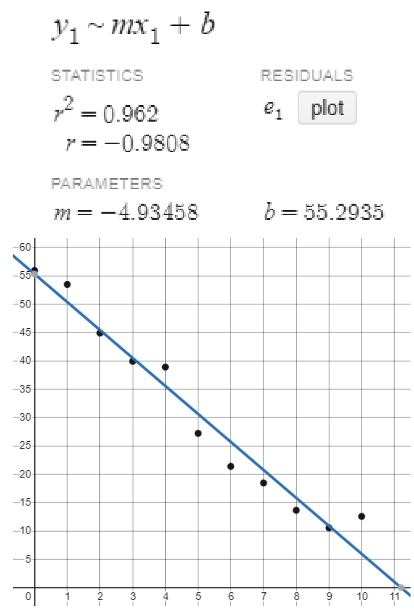
Step 1

- Keep students with their partners.
- Tell students that they will be using technology to find the line of best fit. This process is sometimes referred to as regression, and the line of best fit is represented by a regression equation.
- Demonstrate to students how to use Desmos to calculate the regression equation, display the line of best fit with the scatter plot, and recognize the correlation coefficient. Access the data on the NC New Covid Cases during the 10 weeks after January 3, 2021 from the Activity 1 tab: https://bit.ly/U4L4DataSet.
  - Access the Desmos graphing calculator (www.desmos.com/calculator), click the + icon in the top left corner, and select "table."
  - Enter the data into the table. The points will be graphed, creating the scatter plot. Adjust the graph settings manually or by using the "zoom fit" feature.



Weeks after January 3, 2021	NC new COVID cases (in thousands)
0	55.918
1	53.471
2	44.89
3	39.911
4	38.894
5	27.203
6	21.372
7	18.452
8	13.628
9	10.539
10	12.548

- To calculate the equation for the line of best fit (the regression equation), go to the next line. Type "y1~mx1+b". This will appear as $y_1 \sim mx_1 + b$, as shown.
- The following will be displayed:
  - the statistics: in which the correlation coefficient, r, can be found
  - the parameters of m (the slope) and b (the y-intercept)
  - the graph of the equation displayed with the scatter plot



- Record the regression equation $y = -4.93x + 55.29$ for all to see.
- Ask students to interpret the slope and y-intercept.
  - Provide students with 1 minute of quiet think time and then another minute to share their interpretations with their partner.
- Ask students to share their interpretations. (In the first 10 weeks of 2021, the number of new Covid cases in NC decreased by an average of 4.93 thousand each week. The model estimates that during the week of January 3, 2021, there were 55.29 thousand new cases, which is close to the true value of 55.918 thousand new cases.)

Step 2

- Have students complete questions 1 and 2 and compare their answers with their partner.

Student Task Statement

For each data table, calculate the regression equation and the correlation coefficient. Round values to the nearest hundredth (two decimal places).

1.

x	y
0	3
1	7
2	9
3	10
4	15
5	19
6	21
7	28

2.

x	y
0	28
1	21
2	20
3	19
4	13
5	15
6	11
7	10

Step 3

- Use an abridged *Critique, Correct, Clarify* routine to give students 1–2 minutes in pairs to improve the following statement by making it more clear, more correct, and more complete:

> The correlation coefficient tells you how steep a line is, the slope tells you how well a line fits a data set, and the *y*-intercept tells you the smallest value in the data set.

- Ask two pairs of students to read their improved statement aloud to the class. Scribe the revised statements and invite students to compare the statements with each other and with their own. (For example: "The correlation coefficient tells you how well a line fits a data set, the slope tells you how steep a line is, which represents how fast something is changing, and the y-intercept tells you the point in the data set when $x = 0$.")

DO THE MATH	PLANNING NOTES

Activity 2: Ice Cream Sales (*15 minutes*)

Instructional Routine: Fit It
Addressing: NC.M1.S-ID.6a, NC.M1.S-ID.7

In this *Fit It* activity, students engage in part of the statistical analysis process. They are provided data on average temperature and ice cream sales. The process begins with students creating a scatter plot of the data. Next, students use technology to calculate the regression equation and the correlation coefficient. Students interpret the slope and y-intercept in context and use the regression equation to make predictions. Lastly, students consider the validity of the predictions.

RESPONSIVE STRATEGY

Provide prompts, reminders, or checklists that focus on increasing the length of on-task orientation in the face of distractions. For example, provide two copies of the steps: enter the data in the table, find the best fit line, find the slope, and find the y-intercept. Include phrases to activate knowledge from prior activities such as, "As x increases . . ."

Supports accessibility for: Organization; Conceptual processing; Attention

Step 1

- Keep students with their partners.
- Ask students: "How is the amount of ice cream sold related to the outside temperature?" (The warmer the temperature, the more ice cream sold.)
- Tell students that they will be investigating this using data on average temperature outside and the weight of ice cream sold in pounds. Display the data from the Activity 2 tab: https://bit.ly/U4L4DataSet.
- Ask students to complete questions 1–4.

Monitoring Tip: Monitor for how students make predictions for question 4. Some students may work directly in Desmos, either dragging the cursor along the line of best fit or, for example, graphing $x = 70$ and finding the intersection. Others may use the regression equation and substitute for the known values. Have students share these different strategies during Step 2.

Advancing Student Thinking: Students may struggle with interpreting slope and y-intercept. Remind students of how each relates to a situation. To help students interpret slope, ask them: "What does the x variable represent? What does the y variable represent? What happens to x as y increases (or decreases)? How is slope connected to the x and y variables?" To help students interpret the y-intercept, ask them: "What does the point $(75, 49)$ mean in the scatter plot? What are the coordinates of the y-intercept? What do each of the coordinates mean in the situation described? What is the y value when x is 0? Which variable has a value of 0? Which variable is represented with y?"

Student Task Statement

The average temperature outside in degrees Fahrenheit and the weight of ice cream sold in pounds at a local grocery store are recorded in the table.

Daily high temperature outside in °F (x)	66	69	71	75	77	80	84	90
Weight of ice cream sold in pounds (y)	37	44	43	49	52	61	65	70

1. Use technology to create a scatter plot of this data.
2. Use technology to compute the regression equation.
    a. Record the regression equation. Round values to the nearest hundredth (two decimal places).
    b. Record the correlation coefficient.
3. What are the values for the slope and y-intercept? What do these values mean in this situation?
4. Use the line of best fit to predict:
    a. the number of pounds of ice cream sold if the daily high temperature outside is 70 °F.
    b. the number of pounds of ice cream sold if the daily high temperature outside is 100 °F.
    c. the daily high temperature outside if there were 50 pounds of ice cream sold.

Are You Ready For More?

(*Technology required*.) Priya uses several different ride services to get around her city. The table shows the distance, in miles, she traveled during her last 10 trips and the price of each trip, in dollars.

Distance (miles)	Price (dollars)
3.1	12.5
4.2	14.75
5	16
3.5	13.25
2.5	12
1	9
0.8	8.75
1.6	9.75
4.3	12
3.3	14

1. Priya creates a scatter plot of the data using the distance, x, and the price, y. She determines that a linear model is appropriate to use with the data. Use technology to find the regression equation.
2. Interpret the slope and the y-intercept of the line of best fit in this situation.
3. Use the line of best fit to estimate the cost of a 3.6-mile trip. Will this estimate be close to the actual value? Explain your reasoning.
4. On her next trip, Priya tries a new ride service and travels 3.6 miles, but she pays only $4.00 because she receives a discount. Include this trip in the table and calculate the regression equation for the 11 trips. Did the slope of the line of best fit increase, decrease, or stay the same? Why? Explain your reasoning.
5. Priya uses the new ride service for her 12th trip. She travels 4.1 miles and is charged $24.75. How do you think the slope of the regression equation will change when this 12th trip is added to the table?

Step 2

- Ask selected students to share their responses for the meaning of the slope and *y*-intercept of the line of best fit. (The slope is 1.43 and means there is an increase of 1.43 pounds of ice cream sold for every 1° increase in temperature. The y-intercept is -56.77 and means that at a temperature of 0°, there are -56.77 pounds of ice cream sold. This doesn't make sense in this context and is probably not a good predictor of what really happens at 0°.)

- Instruct students that on question 4a, when predicting the pounds of ice cream sold when the daily high temperature was 70 °F, they were *interpolating* and when predicting the pounds of ice cream sold when the daily high temperature was 100 °F, they were *extrapolating*. Ask students to predict the definitions of “interpolate” and “extrapolate.” (Interpolate is when a predicted value is made from an x-value inside the boundaries of the given data set. Extrapolate is when a predicted value is made from an x-value that falls outside the boundaries of the given data set.)

- Tell students they should be careful when predicting values outside the boundaries of the data set and, in particular, for the y-intercept. Even when the data is fit well by a linear model, the behavior of the variables farther away may not be linear. It is important to remember that all predictions using the line of best fit are estimates and the reasonableness of the predictions should be considered.

DO THE MATH	PLANNING NOTES

Lesson Debrief (*5 minutes*)

The purpose of this lesson is for students to get comfortable calculating a regression equation to model data and then using the equation (and graph) to analyze the data.

Choose what questions to focus the discussion on, whether students should first have an opportunity to reflect in their workbooks or talk through these with a partner, and what questions will be prioritized in the full class discussion.

Display the graph and Desmos output for all to see.

$$y_1 \sim mx_1 + b$$

STATISTICS

$r^2 = 0.9794$

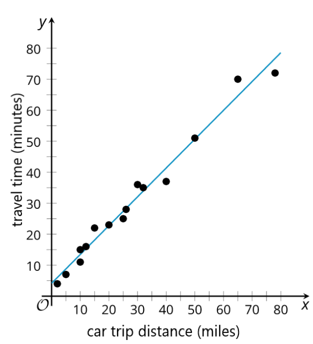
$r = 0.9897$

RESIDUALS

e_1 plot

PARAMETERS

$m = 0.928114$ $b = 4.14614$



Ask students the following questions, prioritizing as necessary:

- "What is the equation of the best fit line? How do we know?" ($y = 0.93x + 4.15$; we use m for the slope and b for the y-intercept). “What is another name for this equation?” (the regression equation)

- "What does the slope of this line mean?" (For every 1-mile increase in the length of a car trip, this linear model predicts that the travel time increases by about 0.93 minutes.)

PLANNING NOTES

- "What does the y-intercept mean?" (It's the predicted time for a trip of length zero miles.) "Is this reasonable?" (No, if a car trip is zero miles, it should not take any time at all.)
- "What is the value of the correlation coefficient?" (approximately 0.99) "What does this mean?" (There is a very strong positive linear relationship between the distance of a car trip and time it takes to get there.)
- "How could we predict the time it will take to make a 41-mile car trip?" (We could substitute 41 for x in the regression equation and calculate y; we could look at the line of best fit and estimate- looks like just over 44 minutes.)

Student Lesson Summary and Glossary

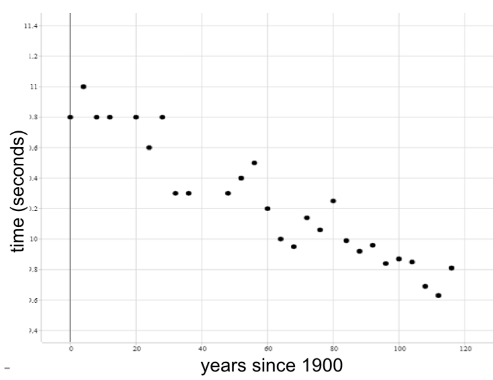
For data sets that appear to have a linear relationship, calculating an equation for the line of best fit can help you describe the relationship between the variables and make predictions. The equation for the line of best fit is called a regression equation.

Take for example the data of winning times for the Men's 100-Meter Run in the Olympics.

Years since 1900	Time (sec)
0	10.8
4	11
8	10.8
12	10.8
20	10.8
24	10.6
28	10.8
32	10.3
36	10.3
48	10.3
52	10.4
56	10.5
60	10.2
64	10

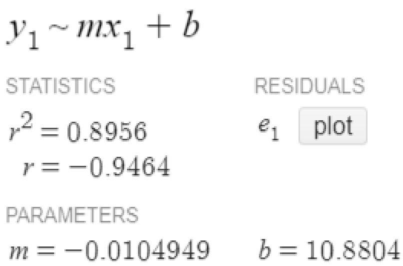
Years since 1900	Time (sec)
68	9.95
72	10.14
76	10.06
80	10.25
84	9.99
88	9.92
92	9.96
96	9.84
100	9.87
104	9.85
108	9.69
112	9.63
116	9.81

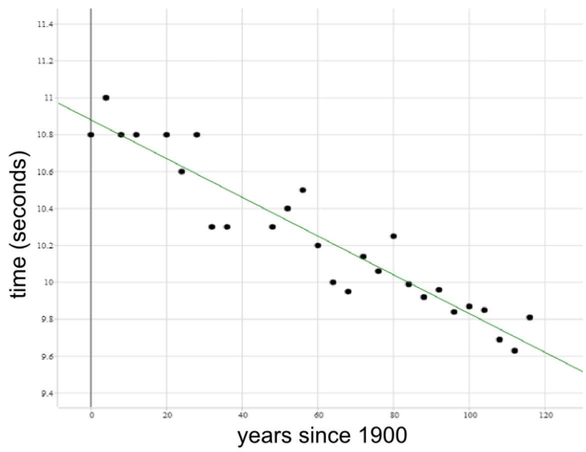
The scatter plot of this data reveals a strong, negative linear relationship.



We can use technology to graph the line of best fit and find the regression equation that describes it. To calculate a line of best fit in Desmos, type "y1~m*x1+b" on the entry line below the table of values. Under "parameters" you will see the slope (m) and y-intercept (b) of the line of best fit. Use these to write the regression equation.

Using this process, we find that the line of best fit has the equation $y = -0.01x + 10.88$. This line is shown below. We can also identify the correlation coefficient as $r = -0.9464$ which indicates a strong, negative relationship between the variables.





The *y*-intercept of this line is 10.88; this is the *predicted* winning time for the Men's 100-Meter Run for the year 1900. This predicted value is very close to the actual winning time of 10.8 seconds. The slope of the best fit line is -0.01. In other words, our linear model predicts that the Men's 100-Meter Run winning time decreases by 0.01 seconds for every year that goes by.

There were no Summer Olympics held in 1940 due to World War II. However, using our model we could predict that the winning time in 1940 would have been:

$y = -0.01(40) + 10.88$

$y = 10.48$ seconds

This is an **interpolated** value and is a reasonable prediction.

Our model can also predict the winning time for the 2024 Olympics:

$y = -0.01(124) + 10.88$

$y = 9.64$ seconds

However, this is an **extrapolated** value. This prediction falls outside the boundaries of our data set and should be interpreted with caution. Why? Imagine extending the line downward and to the right. Eventually, the line will cross the *x*-axis, predicting *negative* times for the Men's 100-Meter Run, which does not make sense. This means that the linear pattern cannot continue forever. Since we don't know at what point the line will stop being a good model, our prediction for the year 2024 is less reliable.

Interpolation: When a predicted value comes from an *x*-value inside the boundaries of the given data set.

Extrapolation: When a predicted value comes from an *x*-value that falls outside the boundaries of the given data set.

Cool-down: Fresh Air (*5 minutes*)

Building On: NC.8.SP.2	Addressing: NC.M1.S-ID.6
Cool-down Guidance: More Chances Students will have more opportunities to understand the mathematical ideas in this cool-down, so there is no need to slow down or add additional work to the next lessons. Instead, use the results of this cool-down to provide guidance for what to look for and emphasize over the next several lessons to support students in advancing their current understanding.	

Provide access to devices that can run Desmos or other graphing technology and this data spreadsheet link: https://bit.ly/U4L4DataSet.

Cool-down

This data can be found https://bit.ly/U4L4DataSet on the Cool-down tab.

The data shows the (*number of trees in a forest, oxygen produced*).

1. Use technology to find the regression equation.
2. Interpret the slope in context. Is the interpretation reasonable? Explain your reasoning.
3. Interpret the y-intercept in context. Is the interpretation reasonable? Explain your reasoning.

Student Reflection: In your own words, how might you describe how to determine if a line is a bad fit, good fit, or best fit? If you are unsure, what is causing confusion for you?

Number of trees in a forest	Tons of oxygen produced by the forest
148	16.43
175	25.64
190	23.28
200	29.2
202	21.41
425	60.56
505	50.75
528	74.45
562	62.66
585	84.24

DO THE MATH

INDIVIDUAL STUDENT DATA

SUMMARY DATA

NEXT STEPS

TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

In the next lesson, students will learn to use residuals to assess the line of best fit. What do you notice in their work from today's lesson that you might leverage in that future lesson?

Practice Problems

1. (*Technology required.*)

x	y
83	102
87	115
91	107
93	122
97	125
97	127
101	120
104	127

a. Use graphing technology to create a scatter plot and find the regression equation for the line of best ft.

b. What does the best fit line estimate for the y value when x is 100?

2. (*Technology required.*)

x	y
2.3	6.2
2.8	5.7
3.1	4.7
3	3.2
3.5	3
3.8	2.8

a. What is the equation of the line of best fit? Round values to the nearest hundredth (two decimal places).

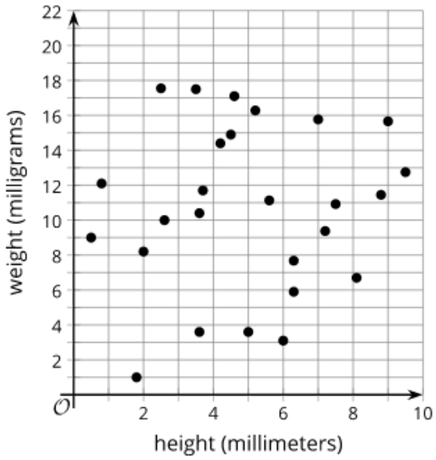
b. What does the equation predict for y when x is 2.3? Round values to the nearest thousandth (three decimal places).

c. How does the predicted value compare to the actual value from the table when x is 2.3?

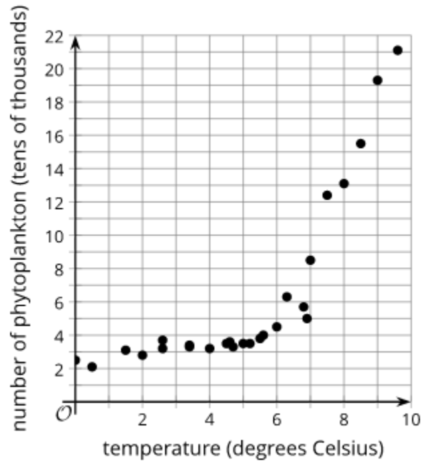
d. How does the predicted value compare to the actual value from the table when x is 3?

3. Which of these scatter plots shows data that would best be modeled with a linear function? Explain your reasoning.

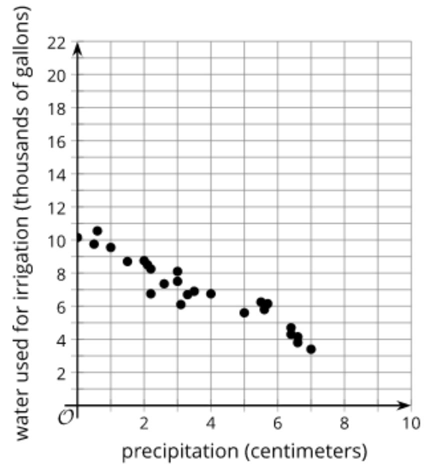
a.



b.



c.



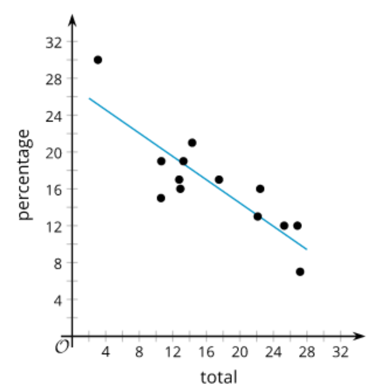
4. A seed is planted in a glass pot, and its height is measured in centimeters every day.

   The best fit line is given by the equation $y = 0.404x - 5.18$, where y represents the height of the plant above ground level and x represents the number of days since it was planted.

   a. What is the slope of the best fit line? What does the slope of the line mean in this situation? Is it reasonable?

   b. What is the y-intercept of the best fit line? What does the y-intercept of the line mean in this situation? Is it reasonable?

   (From Unit 4, Lesson 3)



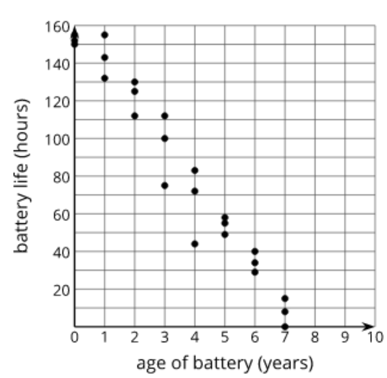
5. At a restaurant, the total bill and the percentage of the bill left as a tip are represented in the scatter plot.

   The best fit line is represented by the equation $y = -0.632x + 27.1$, where x represents the total bill in dollars and y represents the percentage of the bill left as a tip.

   a. What does the best fit line predict for the percentage of the bill left as a tip when the bill is \$15? Is this reasonable?

   b. What does the best fit line predict for the percentage of the bill left as a tip when the bill is \$50? Is this reasonable?

   (From Unit 4, Lesson 3)

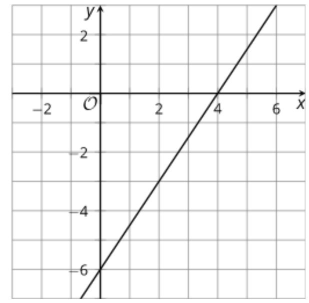


6. A recent study investigated the amount of battery life remaining in alkaline batteries of different ages. The scatter plot shows this relationship between the different alkaline batteries tested.

   The scatter plot includes a point at $(7, 15)$. Describe the meaning of this point in this situation.

   (From Unit 4, Lesson 1)

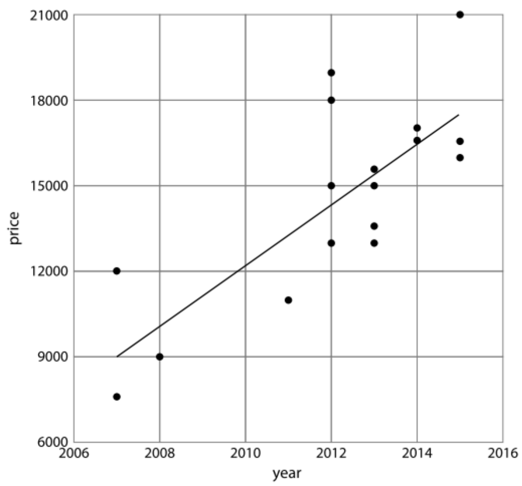
7. Here is the graph of the equation $3x - 2y = 12$. Select **all** the coordinate pairs that represent a solution to the equation.



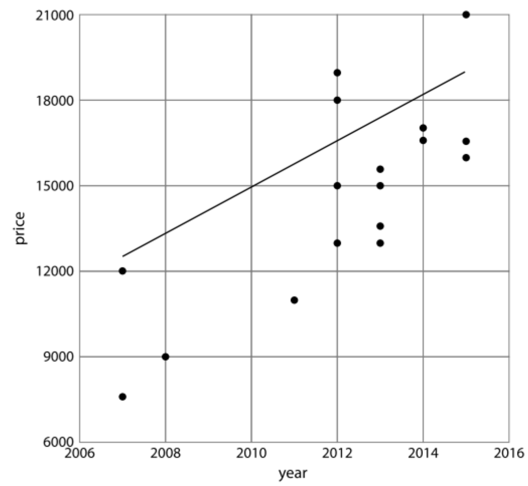
   a. $(2, -3)$
   b. $(4, 0)$
   c. $(5, -1)$
   d. $(0, -6)$
   e. $(2, 3)$

   (Unit 3)

8. 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.[1]



A



B

   a. Look at diagram A.
      - For how many cars does the model in diagram A make a good prediction of its price?
      - For how many cars does the model underestimate the price?
      - For how many cars does it overestimate the price?
   b. Look at diagram B.
      - For how many cars does the model in diagram B make a good prediction of its price?
      - For how many cars does the model underestimate the price?
      - For how many cars does it overestimate the price?
   c. 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?
   d. Which model does a better job of predicting the price of a used car from its year?

   (Addressing NC.8.SP.2)

[1] Adapted from IM 6–8 Math https://curriculum.illustrativemathematics.org/MS/index.html, which was originally developed by Open Up Resources and authored by Illustrative Mathematics, and is copyright 2017–2019 by Open Up Resources. It is licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). OUR's 6–8 Math Curriculum is available at https://openupresources.org/math-curriculum/. Adaptations and updates to IM 6–8 Math are copyright 2019 by Illustrative Mathematics, and are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Lessons 5 & 6: Checkpoint

PREPARATION

Lesson Goals	Learning Targets
• Learn and grow mathematically in course-level content. • Communicate and address mathematical areas of strength and areas of growth.	• I can continue to grow as a mathematician and challenge myself. • I can share what I know mathematically.

Lesson Narrative

This is a Checkpoint day. These lessons have three main purposes: 1. differentiated and small-group instruction; 2. opportunities for students to participate in various learning stations to refine and extend previous learning; and 3. the opportunity for students to complete the next unit's Check Your Readiness (CYR). Administering the CYR at this point in the unit allows plenty of time for the data to inform the next unit's instruction. Checkpoint days consist of two lessons (one full block) and are structured as four 20-minute stations that students rotate between. There are a total of seven stations students can engage with. Since students will not be able to participate in all seven stations, please note that Station A (Unit 5 Check Your Readiness) is required for all students.

A. Unit 5 Check Your Readiness (*Required*)
B. Teacher-led Small-group Instruction
C. Laptop Battery Charge
D. School Data
E. Are You Ready For More?
F. A Midpoint Miracle
G. Cat and Dog Food

What do your students think it means to be good at math? How are you helping them change negative impressions they might have about their ability to reason mathematically?

When do your students feel successful in math? How do you know?

Agenda, Materials, and Preparation

- **Station A** (*Required, 20 minutes*)
  - Unit 5 Check Your Readiness (print 1 copy per student)
- **Station B** (*20 minutes*)
- **Station C** (*20 minutes*)
- **Station D** (*20 minutes*)
  - North Carolina School Report Cards website link: https://ncreports.ondemand.sas.com/src/ (or other relevant school data websites)
- **Station E** (*20 minutes*)
  - Are You Ready For More? tasks in Student Workbook from past lessons (or optional: print 1 blackline master per student)
- **Station F** (*20 minutes*)
- **Station G** (*20 minutes*)

STATIONS

Station A: Unit 5 Check Your Readiness (*Required, 20 minutes*)

Remind students that it is really important that their responses to these questions accurately represent what they know. Ask them to answer what they can to the best of their ability. If they get stuck, they should name what they don't know or understand.

Station B: Teacher-led Small-group Instruction (*20 minutes*)

Use student cool-down data, Check Your Readiness Unit 4 data, informal formative assessment data from Unit 4 (Lessons 1–4), or topics from prior units to provide targeted small-group instruction to students who demonstrate the need for further support or challenge.

Potential topics:

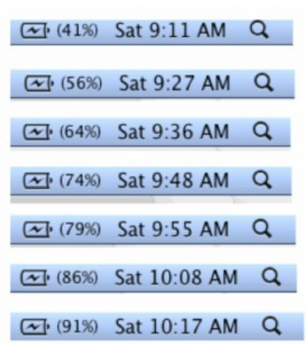
- Scatter plots and linear models
- Revisiting graphing, solving systems of equations and inequalities, or coordinate geometry

Station C: Laptop Battery Charge[1] (*20 minutes*)

Building On: NC.8.SP.2	**Addressing:** NC.M1.S-ID.6

Station C

Elena forgot to plug in her laptop before she went to bed. She wants to take the laptop to her friend's house with a full battery. The pictures below show screenshots of the battery charge indicator after she plugs in the computer at 9:11 a.m.



1. Use the screenshots to consider the questions below.
   a. The screenshots suggest a relationship between two variables. What are the two variables in this situation?
   b. Use technology to make a scatter plot of the data.
   c. Use technology to draw a line that fits the data and find the equation of the line in slope-intercept form.
   d. What is the slope of the line? Explain the meaning of the slope in the context of the problem.

[1] Adapted from https://tasks.illustrativemathematics.org/

e. What is the y-intercept of the equation of the line? Explain the meaning of the y-intercept in the context of the problem.

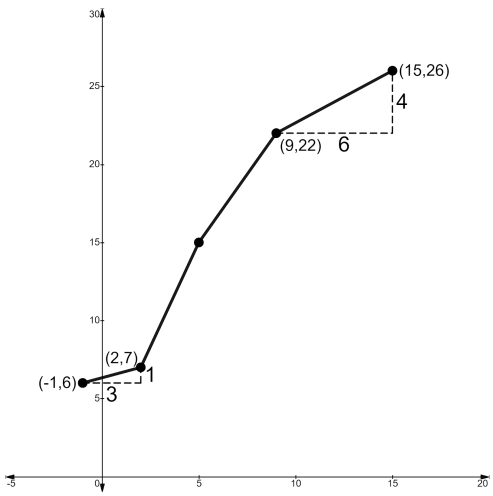
f. At what time does your model predict the battery was 50% charged?

g. When can Elena expect to have a fully charged battery?

2. At 9:27 AM Elena makes a quick calculation: *The battery seems to be charging at a rate of 1 percentage point per minute. So the battery should be fully charged at 10:11 AM.* Explain Elena's calculation. Is her estimate most likely an under- or over-estimate? How does it compare to your prediction?

3. Compare the **average rate of change** of the battery charging function on the first given time interval and on the last given time interval. What does this tell you about how the battery is charging?

Learn More



Average rate of change:

Recall that rate of change is how one quantity changes in relation to another quantity.

Linear functions will have a constant rate of change, while non-linear functions' rates of change vary based on the interval of the graph you are looking at.

In the non-linear graph to the left, for example, the rate of change between $(-1,6)$ and $(2,7)$ is $\frac{7-6}{2-(-1)}$, or $\frac{1}{3}$.

The rate of change between $(9,22)$ and $(15,26)$ is $\frac{26-22}{15-9}$, or $\frac{4}{6}$, which is $\frac{2}{3}$.

A sample comparison of the average rate of change of the first and last intervals might look like:

> The first interval has an average rate of change of $\frac{1}{3}$, while the last interval has an average rate of change of $\frac{2}{3}$. The rate of change of the last interval is greater than the rate of change of the first interval meaning the graph is steeper in the last interval compared to the first.

4. How long would it take for the battery to charge if it started out completely empty?

DO THE MATH

PLANNING NOTES

Station D: School Data (*20 minutes*)

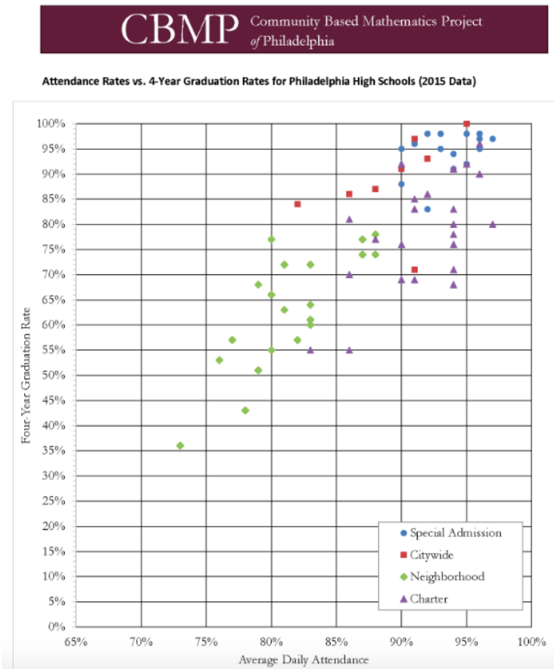
Instructional Routine: Aspects of Mathematical Modeling

Station D offers students an opportunity to engage in *Aspects of Mathematical Modeling* as they research characteristics of high schools in their district or state. Research links to share with students to access the necessary information. For example, in North Carolina, share the following link with students: https://ncreports.ondemand.sas.com/src/. Determine whether students will choose their own set of 10–20 schools or if it would be more beneficial to create a bank of schools for students to choose from.

Station D

1. In Philadelphia, PA there are four types of high schools: neighborhood schools, citywide schools, charter schools, and special admission schools. The graph below compares the average daily attendance rates at these different schools with the rate of students who graduate from those schools in four years.

   What do you notice? What do you wonder?



2. Consider the questions below.

   a. Choose 10–20 high schools in your district or state and research two characteristics you are interested in. Here are some examples:

      - four-year graduation rate compared to percentage of chronic absenteeism
      - class size for Math 1 compared to four-year graduation rate
      - percentage of students participating in an AP assessment compared to percentage of students attending college after graduation
      - percentage of economically disadvantaged students compared to number of suspensions or expulsions per 1000 students

   b. Do you predict any trends?

   c. Create a scatter plot of the data and determine the line of best fit. Is there a relationship between the two characteristics you researched? What trends do you see? What follow-up questions or relationships does this make you want to research? What might be some possible reasons for these trends?

DO THE MATH

PLANNING NOTES

Station E: Are You Ready For More? (*20 minutes*)

Students who did not complete the "Are You Ready For More?" task statements from Lessons 1, 3, and 4 can do so in Station E. This is a great opportunity for students to expand their thinking. These tasks can also be offered as additional practice problems or used in the teacher-led small-group instruction.

Station E

1. Students in Charlotte, NC were interested in examining the access to organic produce in different parts of their city. They collected the following data. In this case, they also collected the population within the neighborhood (defined by zip code).

Population	Median household income (2019)	Organic produce available
71048	65963	27
59664	93942	40
49635	59438	43
9280	136333	44
53629	51676	44
37286	91494	44
37309	45808	46
11315	88039	47

Population	Median household income (2019)	Organic produce available
11195	92786	55
43931	52766	55
42263	71914	55
19283	93938	56
28523	90057	57
20317	76022	58
47208	49465	59

   a. Create a scatter plot for the (*median household income, organic produce available*) and describe any relationship between the two variables.

   b. Compare this relationship to the one you found for San Antonio. What do you think are the reasons for any similarities or differences?

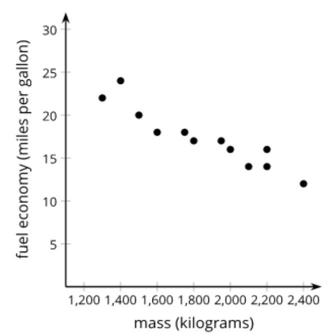
   c. Create a scatter plot for the (*population, organic produce available*) and describe any relationship between the two variables.

   d. One of the points appears to be an outlier. How does your answer to question 3 change if the outlier is removed?

   (From Unit 4, Lesson 1)

2. Clare, Diego, and Elena collect data on the mass and fuel economy of cars at different dealerships. Clare finds the line of best fit for data she collected for 12 used cars at a used car dealership. The line of best fit is $y = \frac{-9}{1000}x + 34.3$, where x is the car's mass, in kilograms, and y is the fuel economy, in miles per gallon.

   Diego made a scatter plot for the data he collected for 10 new cars at a different dealership.



   Elena made a table for data she collected on 11 hybrid cars at another dealership.

Mass (kilograms)	Fuel economy (miles per gallon)
1,100	38
1,200	39
1,250	35
1,300	36
1,400	31
1,600	27
1,650	28
1,700	26
1,800	28
2,000	24
2,050	22

   a. Interpret the slope and y-intercept of Clare's line of best fit in this situation.

   b. Diego looks at the data for new cars and used cars. He claims that the fuel economy of new cars decreases as the mass increases. He also claims that the fuel economy of used cars increases as the mass increases. Do you agree with Diego's claims? Explain your reasoning.

   c. Elena looks at the data for hybrid cars and correctly claims that the fuel economy decreases as the mass increases. How could Elena compare this decrease for hybrid cars to the decrease for new cars? Explain your reasoning.

   (From Unit 4, Lesson 3)

3. (*Technology required*.) Priya uses several different ride services to get around her city. The table shows the distance, in miles, she traveled during her last 10 trips and the price of each trip, in dollars.

Distance (miles)	Price (dollars)
3.1	12.5
4.2	14.75
5	16
3.5	13.25
2.5	12
1	9
0.8	8.75
1.6	9.75
4.3	12
3.3	14

   a. Priya creates a scatter plot of the data using the distance, x, and the price, y. She determines that a linear model is appropriate to use with the data. Use technology to find the regression equation.

   b. Interpret the slope and the y-intercept of the line of best fit in this situation.

   c. Use the line of best fit to estimate the cost of a 3.6-mile trip. Will this estimate be close to the actual value? Explain your reasoning.

   d. On her next trip, Priya tries a new ride service and travels 3.6 miles, but she pays only $4.00 because she receives a discount. Include this trip in the table and calculate the regression equation for the 11 trips. Did the slope of the line of best fit increase, decrease, or stay the same? Why? Explain your reasoning.

   e. Priya uses the new ride service for her 12th trip. She travels 4.1 miles and is charged $24.75. How do you think the slope of the regression equation will change when this 12th trip is added to the table?

(From Unit 4, Lesson 4)

DO THE MATH

PLANNING NOTES

Station F: A Midpoint Miracle[2] (*20 minutes*)

This task, revisiting a topic from Unit 3, gives students the opportunity to prove a surprising fact about quadrilaterals: that if the midpoints of an arbitrary quadrilateral are joined to form a new quadrilateral, then the new quadrilateral is a parallelogram, even if the original quadrilateral was not.

One option is to have students draw the quadrilateral in Desmos. This allows students to move various parts of the quadrilateral around and see that the midpoint quadrilateral remains a parallelogram no matter where the parts are moved. This strategy also allows students to see that in general, the midpoint quadrilateral is not a rhombus or rectangle.

Addressing: NC.M1.G-GPE.4; NC.M1.G-GPE.5

[2] Adapted from https://tasks.illustrativemathematics.org/

Station F

Let's investigate quadrilaterals further by examining midpoint quadrilaterals!

1. Construct a quadrilateral and its midpoint quadrilateral using the following steps:

   Step 1: Go to desmos.com/geometry.

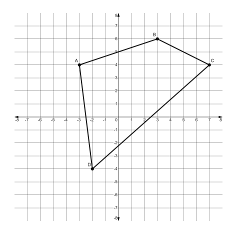
   Step 2: In the "Construct" panel on the left, select "Polygon." Click on one point on the grid and then click on another point, creating the first side of your quadrilateral. Click on two more grid points to make the other two vertices of the quadrilateral. End by connecting the 4th side to the very first point you clicked to close your quadrilateral.

   Step 3: Create a midpoint quadrilateral:

   - Click on "Midpoint" under "More Tools" in the "Construct" panel. Then, click on one side of the quadrilateral. Desmos will mark the midpoint of that side. Repeat this step until midpoints of each side of the quadrilateral have been located.
   - Connect the midpoints to form a midpoint quadrilateral by selecting the "Polygon" tool in the "Construct" panel and clicking on each midpoint, going clockwise, ending by clicking on the midpoint you started with.

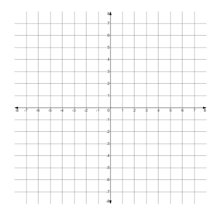
2. What do you notice about the midpoint quadrilateral?

3. Create different versions of the original quadrilateral by dragging its vertices around using the "Select" tool in the "Construct" panel. As you drag points, the midpoint quadrilateral will adjust accordingly. Does your observation in Question 2 seem to be true for all of the midpoint quadrilaterals? If you have a different observation now, what is it?



4. In this diagram, $ABCD$ is the original quadrilateral. Find the midpoint of each side, and label the midpoints *P*, *Q*, *R*, and *S*. Draw the midpoint quadrilateral $PQRS$. Now prove that your observation from Questions 2 or 3 is also true for quadrilateral $PQRS$.

5. If time allows, and you are ready for more, create your own quadrilateral $ABCD$ (and corresponding midpoint quadrilateral $PQRS$) on the coordinate plane below to further test your claim.



DO THE MATH

PLANNING NOTES

Station G: Cat and Dog Food (*20 minutes*)

In Station G, students are presented with a series of problems that can be modeled by systems of equations. Some problems can be solved simply by reasoning; others more naturally prompt writing a system. This station can be done independently, with a partner, or in the teacher-led small group.

Building On: NC.8.EE.8	**Addressing:** NC.M1.A-CED.3; NC.M1.A-REI.5; NC.M1.A-REI.6

Station G

Andre works at Pet Palace grooming and kennel. He is in charge of purchasing the food and other supplies for the cats and dogs to use during their stay. Andre's manager asked for the cost for dog food, cat food, leashes, and brushes, but Andre only has information for the total amount he spent. Use the information below to calculate the costs for each of these items.[3]

1. One week, Andre bought three bags of *Tabitha Tidbits* and four bags of *Figaro Flakes* for $43.00. The next week, he bought three bags of *Tabitha Tidbits* and six bags of *Figaro Flakes* for $54.00. Based on this information, figure out the price of one bag of each type of cat food. Explain your reasoning.

2. One week, Andre bought two bags of *Brutus Bites* and three bags of *Milo Munchies* for $42.50. The next week, he bought five bags of *Brutus Bites* and six bags of *Milo Munchies* for $94.25. Based on this information, figure out the price of one bag of each type of dog food. Explain your reasoning.

3. Andre purchased six dog leashes and six cat brushes for $45.00 for Elena to use while pampering the pets. Later in the summer, he purchased three additional dog leashes and two cat brushes for $19.00. Based on this information, figure out the price of each item. Explain your reasoning.

4. One week Andre purchased two boxes of cat treats and three boxes of dog treats for $18.50. The next week, he bought two boxes of cat treats and two boxes of dog treats for $14.00. The third week, he bought five boxes of both cat and dog treats for $35.00. Based on this information, figure out the price of each item. Explain your reasoning.

5. Andre has noticed that because his purchases have been somewhat similar, it has been easy to figure out the cost of each item. However, his last set of receipts has him puzzled. One week, he tried out cheaper brands of cat and dog food. On Monday, he purchased three small bags of cat food and five small bags of dog food for $22.75. Because he went through the small bags quite quickly, he had to return to the store on Thursday to buy two more small bags of cat food and three more small bags of dog food, which cost him $14.25. Based on this information, figure out the price of each bag of the cheaper dog food. Explain your reasoning.

DO THE MATH

PLANNING NOTES

[3] Adapted from Secondary Math 1, Module 5, Mathematics Vision project http://www.mathematicsvisionproject.org, licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0)

TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

Reflect on how comfortable your students are asking questions of you and of each other. What can you do to encourage students to ask questions?

Lesson 7: Residuals

PREPARATION

Lesson Goals	Learning Target
● Calculate and plot the residuals for a given data set and use the information to determine the goodness of fit for a linear model. ● Comprehend the connection between residuals, variability, and whether or not using a linear model is appropriate.	● I can plot and calculate residuals for a data set and use the information to judge whether a linear model is a good fit.

Lesson Narrative

The mathematical purpose of this lesson is to informally assess the fit of a function by plotting and analyzing residuals. The term **residual** is introduced as the difference between the y-value for a point in a scatter plot and the value predicted by the linear model for the associated x-value. The work of this lesson connects to previous work because students analyzed bivariate data by creating scatter plots, used the correlation coefficient to determine appropriateness of a linear model, and fit linear functions to the data. In the next lesson, students will analyze these relationships to determine if the correlation of the linear model is causal or associative.

When students take turns with a partner matching graphs of residuals to scatter plots that display linear models, students trade roles explaining their thinking and listening, providing opportunities to explain their reasoning and critique the reasoning of others (MP3).

Where do you see connections from what students shared and discussed in previous lessons to this lesson?

Focus and Coherence

Building On	Addressing
NC.7.NS.1: Apply and extend previous understandings of addition and subtraction to add and subtract rational numbers, using the properties of operations, and describing real-world contexts using sums and differences. **NC.8.SP.2:** Model the relationship between bivariate quantitative data to: ● Informally fit a straight line for a scatter plot that suggests a linear association. ● Informally assess the model fit by judging the closeness of the data points to the line.	**NC.M1.S-ID.6:** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related. **a.** Fit a least squares regression line to linear data using technology. Use the fitted function to solve problems. **b.** Assess the fit of a linear function by analyzing residuals. **NC.M1.S-ID.8:** Analyze patterns and describe relationships between two variables in context. Using technology, determine the correlation coefficient of bivariate data and interpret it as a measure of the strength and direction of a linear relationship. Use a scatter plot, correlation coefficient, and a residual plot to determine the appropriateness of using a linear function to model a relationship between two variables.

Agenda, Materials, and Preparation

A data spreadsheet will be used in Activity 1 and for Practice Problem #5. Students can access this data through this link: https://bit.ly/U4L7DataSet.

- **Bridge** (*Optional, 5 minutes*)
- **Warm-up** (*5 minutes*)
- **Activity 1** (*15 minutes*)
  - Technology is required for this activity: Acquire devices that can access Desmos (recommended) or other graphing technology. It is ideal if each student has their own device.
- **Activity 2** (*10 minutes*)
  - Best Residuals card sort (print page 1 per pair of students and cut up in advance; print page 2 for the teacher demonstration set only)
- **Lesson Debrief** (*5 minutes*)
- **Cool-down** (*5 minutes*)
  - M1.U4.L7 Cool-down (print 1 copy per student)

LESSON

Bridge (*Optional, 5 minutes*)

Building On: NC.8.SP.2

The purpose of this bridge is to provide students the opportunity to refine and consolidate their understanding of when a good linear model exists for a data set: namely, a line can be drawn so that the data points are scattered fairly evenly above and below the line, and most points are close to the line. Students will expand on this concept later in this lesson by assessing the fit of a linear function by analyzing residuals. This task is aligned to questions 2 and 3 in Check Your Readiness.

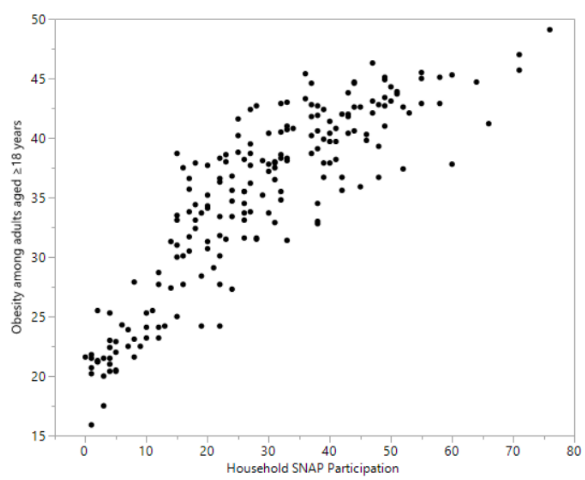
Student Task Statement

The following scatter plot represents the frequency of obesity based on household participation in SNAP (food stamps) by census tract in Baltimore, MD.[1]

a. Draw a linear model for the data, paying attention to the closeness of the data points on either side of the line.

b. Estimate the equation of the best fit line using the slope and y-intercept of your linear model.

c. Lin wanted to read more about this and learned that there are many factors that contribute to this relationship that include affordability of healthy food, access to quality healthcare, and access to safe and affordable places to exercise. In what way could access to grocery stores with healthy options also play a role in this relationship?



(Image source[2])

[1] North Carolina Department of Public Instruction. (2018). *Instructional Support Tools for Achieving New Standards: 8th Grade.* https://files.nc.gov/dpi/documents/curriculum/mathematics/scos/current/8th-unpacking.pdf

[2] Borders, S. (2018). *Barriers to and Facilitators of, Providing an Effective Approach to Preventing Obesity in Women Ages 18 to 39 A Geospatial Analysis of Baltimore.* https://www.researchgate.net/publication/328489797_Barriers_to_and_Facilitators_of_Providing_an_Effective_Approach_to_Preventing_Obesity_in_Women_Ages_18_to_39_A_Geospatial_Analysis_of_Baltimore_Barriers_to_and_Facilitators_of_Providing_an_Effective_A

DO THE MATH	PLANNING NOTES

Warm-up: Differences in Expectations (*5 minutes*)

Instructional Routines: Math Talk; Discussion Supports (MLR8) - Responsive Strategy	
Building On: NC.7.NS.1	**Building Towards:** NC.M1.S-ID.6b

The purpose of this *Math Talk* is to elicit strategies and understandings students have for subtracting a predicted value from an actual value. These understandings help students develop fluency and will be helpful later in this lesson when students will need to be able to compute residuals from a linear model.

Step 1

- Before beginning the Math Talk, display the following:

> Mentally calculate how close the predicted value is to the actual value using the difference: $actual\ value - predicted\ value$.
>
> – Actual value: 21 cars. Predicted value: 20 cars
>
> – Actual value: 20 cars. Predicted value: 21 cars

RESPONSIVE STRATEGY

To support working memory, provide students with sticky notes or mini whiteboards.

Supports accessibility for: Memory; Organization

- Ask students, “How would the $actual\ value - predicted\ value$ be different for the two situations? (1 car vs -1 car)
- Display one problem at a time. Give students quiet think time for each problem and ask them to give a signal when they have an answer and a strategy.
- Keep all problems displayed throughout the talk.

Student Task Statement

Mentally calculate how close the predicted value is to the actual value using the difference: $actual\ value - predicted\ value$.

1. Actual value: 24.8 grams. Predicted value: 19.6 grams
2. Actual value: $112.11. Predicted value: $109.30
3. Actual value: 41.5 centimeters. Predicted value: 45.90 centimeters
4. Actual value: -1.34 degrees Celsius. Predicted value: -2.45 degrees Celsius

Step 2

- Facilitate a whole-class discussion by asking students to share their strategies for each problem. Record and display their responses for all to see. To involve more students in the conversation, consider asking:
  - “Who can restate ___’s reasoning in a different way?”
  - “Did anyone have the same strategy but would explain it differently?”
  - “Did anyone solve the problem in a different way?”
  - “Does anyone want to add on to ___’s strategy?”
  - “Do you agree or disagree? Why?”

RESPONSIVE STRATEGY

Provide sentence frames to support students when they explain their strategy. For example: “First, I ___ because . . . .” or “I noticed ___ so I . . . .” Some students may benefit from the opportunity to rehearse what they will say with a partner before they share with the whole class.

Discussion Supports (MLR8)

DO THE MATH	PLANNING NOTES

Activity 1: Oranges Return (*15 minutes*)

Instructional Routines: Fit It; Collect and Display (MLR2)
Addressing: NC.M1.S-ID.6a, b, NC.M1.S-ID.8

This *Fit It* activity uses the data from the video of orange weights from the Orange You Glad We’re Boxing Fruit activity in Lesson 3. This activity introduces the concept of residuals and has students plot and analyze the residuals to informally assess the fit of a function. In this activity, students also fit a function to data, and use the function to solve problems. Students learn that a residual is the difference between the actual y-value for a point and the predicted y-value for the point on the linear model with the same associated x-value.

Step 1

- Ask students to arrange themselves in pairs or use visibly random grouping.
- Display the data from the video about weighing oranges as well as the photograph of the oranges on the scale for all to see.

Number of oranges	Weight in kilograms
3	1.027
4	1.162
5	1.502
6	1.617
7	1.761
8	2.115
9	2.233
10	2.569

- Inform students that for this activity, they will be working with graphing technology and will use their partner to assist them with any technological challenges.
- Encourage students to work independently on questions 1–3 (with technology assistance from their partner, as needed) and then with their partner to explain or show their reasoning for questions 4 and 5. They should, again, work independently on the remaining questions, and then share their thinking with their partner for questions 8–9.
  - Using the *Collect and Display* routine, listen for and collect words and phrases students use to describe how to calculate residuals, what positive residual and negative residual mean, and what it means for the residual to be zero. Display words and phrases such as "difference between actual data and predictions," "prediction is greater/less than actual data," and "linear prediction is close to actual data" for all to see, and then encourage students to use this language in their written responses. This will help students read and use mathematical language during whole-group discussions.

RESPONSIVE STRATEGY

Use color coding and annotations to highlight connections between representations in a problem. For example, highlight the linear column in the table the same color as the line on the graph, and highlight the actual values on the table the same color as the corresponding points on the graph. To extend this further, the residual values can be displayed and highlighted in the same color as the residual points on the graph itself.

Supports accessibility for: Visual-spatial processing

Advancing Student Thinking: Students may not understand how to determine if the linear model predicts the weight of oranges well or poorly. Ask them to determine the weight that the model predicts, then ask how close that prediction is to the actual weight.

Since the residuals will be close to zero, with some negative and some positive, students may need to adjust their graphing windows to see the residuals plot.

Student Task Statement

1. Use technology to make a scatter plot of orange weights and find the line of best fit.

   a. In desmos.com/calculator, click on the + icon to add a table. In the table, enter in the following values, which can also be found on the Activity 1 tab: https://bit.ly/U4L7DataSet.

   b. Find the line of best fit for this scatter plot. Under the table, type the following in a new entry line: $y_1 \sim m \cdot x_1 + b$

Number of oranges	Weight in kilograms
3	1.027
4	1.162
5	1.502
6	1.617
7	1.761
8	2.115
9	2.233
10	2.569

Use the following table for questions 2–5.

Number of oranges	Weight in kilograms	Linear prediction weight in kilograms
3	1.027	
4	1.162	
5	1.502	
6	1.617	
7	1.761	
8	2.115	
9	2.233	
10	2.569	

2. What does the linear model predict for the weight of the box of oranges for each of the number of oranges? Complete the "linear prediction weight in kilograms" column. (To show the y-value predicted by the linear model at a given x-value, click and hold on the best fit line at the x-value you're considering.)
3. Compare the actual weight of the box with three oranges in it to the predicted weight of the box with three oranges in it. Explain or show your reasoning.
4. How many oranges are in the box when the linear model predicts the weight best? Explain or show your reasoning.
5. How many oranges are in the box when the linear model predicts the weight least well? Explain or show your reasoning.

6. The difference between the actual value and the value predicted by a linear model is called the residual. If the actual value is greater than the predicted value, the residual is positive. If the actual value is less than the predicted value, the residual is negative. For the orange weight data set, what is the residual for the best fit line when there are 3 oranges? In Desmos, on the same axes as the scatter plot, plot this residual point using the format (3, residual value) on the next entry line.

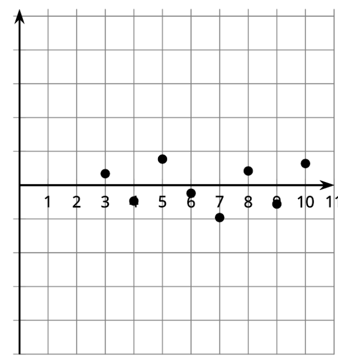
7. In Desmos, it is possible to graph the residuals all at once by clicking the button under residuals labeled "plot." The residual values are also automatically added to your scatterplot table. Check out the graph of the residuals. What is the residual value for eight oranges and what are the coordinates of the point on the residual plot?

RESIDUALS
e_1 plot

8. Which point on the scatter plot has the residual closest to zero? What does this mean about the weight of the box with that many oranges in it?

9. How can you use the residuals to decide how well a line fits the data?

Step 2



- Compare student answers to the question about the point that the line predicts best (question 4) to the answer for the question about the residual closest to zero (question 8).
- Show a graph of the residuals.
- Facilitate a whole-class discussion, referring to any collected and displayed student language. Supplement and update the displayed student language with any clarifications or details as students discuss the following questions:
  - "What does it mean for the residual to be positive? Negative?" (The residual is positive when the actual data value is greater than what the model predicts for that x value and negative when the actual data value is less than the prediction.)
  - "What does it mean when a residual is on or close to the horizontal axis?" (It means that the line of best fit passes through or comes close to passing through that point in the graph.)
  - "Find the residual that has the furthest vertical distance from the horizontal axis. What does this mean in the context of the scatter plot and the line of best fit?" (The residual that is furthest from the horizontal axis has the same x-coordinate as the point that is the greatest vertical distance away from the line of best fit in the scatter plot.)

DO THE MATH

PLANNING NOTES

Activity 2: Best Residuals (*10 minutes*)

Instructional Routines: Card Sort; Stronger and Clearer Each Time (MLR1)	
Addressing: NC.M1.S-ID.6b; NC.M1.S-ID.8	

In this *Card Sort* activity, students work with a partner, matching graphs of residuals to scatter plots that display linear models. As students work together, they alternate explaining their thinking and listening, providing opportunities to explain their reasoning and critique the reasoning of others (MP3). They should begin to recognize that a plot of the residuals for data that is fit well by a linear model shows residuals that are close to the x-axis and do not show a noticeable trend.

Step 1

- Keep students in the same pairs and give each group a set of cut-up Best Residual cards for matching using the *Card Sort* routine.
- Demonstrate how to set up and find matches. Choose a student to be your partner. Mix up the cards and place them face up. Point out that the cards contain either a scatter plot with a linear model or a graph of the residuals. Select the cards labeled "For Teacher Demonstration" and explain to your partner why you think the cards match. Demonstrate productive ways to agree or disagree—for example, by explaining your mathematical thinking or asking clarifying questions.

Step 2

- Use the *Stronger and Clearer Each Time* routine to provide students with multiple opportunities to clarify their explanations through conversation. Give students time to meet with two partners to share their response to the last question.
- Students should first check to see if they agree with each other about who had the best estimate of the line of best fit. Then partners should take turns explaining the reasoning for their choice and providing each other with feedback on their explanations.
- Provide listeners with prompts for feedback that will help their partner add detail to strengthen and clarify their ideas. For example, students can ask their partner: "How did you use the graph of the residuals to make your decision?" or "How do you know . . .?"
- Have students repeat the process with another partner.
- Finally, provide students with 2–3 minutes to revise their initial draft based on feedback from their peers. This will help students produce a written justification for determining best estimates of lines of best fit.

RESPONSIVE STRATEGIES

Represent the same information through different modalities by showing the relationship between the original graph and the graph of the residuals. Provide students with tracing paper (to use directly over the card) or graph paper (to recreate the card) so they can draw in the distance between the point and the line to see the residuals.

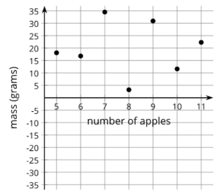
Supports accessibility for: Conceptual processing; Visual-spatial processing

Student Task Statement

1. Match the scatter plots and given linear models to the graph of the residuals.
2. Turn the scatter plots over so that only the residuals are visible. Based on the residuals, which line would produce the most accurate predictions? Which line fits its data worst?

Are You Ready For More?

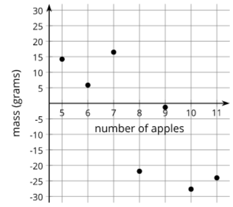
1. Tyler estimates a line of best fit for some linear data about the mass, in grams, of different numbers of apples. Here is the graph of the residuals.



    a. How do the points on the scatter plot compare to Tyler's best fit line?

    b. How well does Tyler's line of best fit model the data? Explain your reasoning.

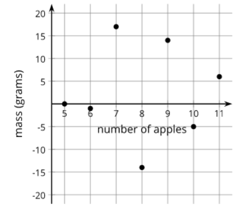
2. Lin estimates a line of best fit for the same data. The graph shows the residuals.



    a. How do the points on the scatter plot compare to Lin's best fit line?

    b. How well does Lin's line of best fit model the data? Explain your reasoning.

3. Kiran also estimates a line of best fit for the same data. The graph shows the residuals.



    a. How do the points on the scatter plot compare to Kiran's best fit line?

    b. How well does Kiran's line of best fit model the data? Explain your reasoning.

4. Who has the best estimate of the line of best fit—Tyler, Lin, or Kiran? Explain your reasoning.

Step 3

- Much discussion takes place between partners. Invite students to share how they made the matches with the whole class.
- Listen for and amplify responses that describe measuring the distance between each point on the scatter plot and the line of best fit and connecting points above or below the line of best fit to positive or negative residual values.

DO THE MATH	PLANNING NOTES

Lesson Debrief (*5 minutes*)

The purpose of this lesson is to understand the connections between a scatter plot displaying a linear model and a graph of the residuals. A good linear model for the data will have residuals that are scattered on either side of the x-axis without a clear pattern and close to the axis.

Choose what questions to focus the discussion on, whether students should first have an opportunity to reflect in their workbooks or talk through these with a partner, and what questions will be prioritized in the full class discussion. Discuss questions such as: • "Reflecting on the card sort, what do you notice about the residuals in graph 2? Explain what you notice in the context of scatter plot A." (The residuals go in a u-shaped pattern. The data set graphed in the scatter plot does not appear linear and is curved, so when you plot a line through it, you would expect the residuals to show the curvature.) • "Describe any difficulties you experienced with the card sort activity and how you resolved them." (It was tough figuring out how to decide where to begin when finding matches. I used my partner's strategy of looking at the values on the x-axis to help narrow down the choices.) • "When looking at the residuals for a linear model for data following a linear trend, Priya found that roughly half of the residuals were positive and the other half were negative. Do her findings about the residuals provide evidence to support the claim that the linear model used is a line of best fit? What else should Priya look for?" (Yes, there is evidence to support the claim. A line of best fit should go through the middle of the data, so roughly half the points should be above the line of best fit and the other half below the line of best fit. Priya should also look for any patterns in the residuals. If there is a pattern, then it is likely not the line of best fit.) • "How can you use a graph of the residuals to informally assess the fit of a function?" (You look to see how far away the residuals are from the horizontal axis. The closer they are, the better the fit. Also, look to see if the positive and negative residuals are distributed randomly. If most of them are positive or negative, or if they form some pattern, then the function is probably not a great fit.)	**PLANNING NOTES**

Student Lesson Summary and Glossary

When fitting a linear model to data, it can be useful to look at the residuals. **Residuals** are the difference between the y-value for a point in a scatter plot and the value predicted by the linear model for that x-value.

> **Residual**: The difference between the y-value for a point in a scatter plot and the value predicted by the linear model for the associated x-value.

In the scatter plot showing the length of fish and the age of fish, the point (2,100) represents a fish who is 2 years old and 100 mm long. We can compare this to what the best fit line predicts for the length of a two-year-old-fish:

$$y = 34.08 \cdot 2 + 23.78$$

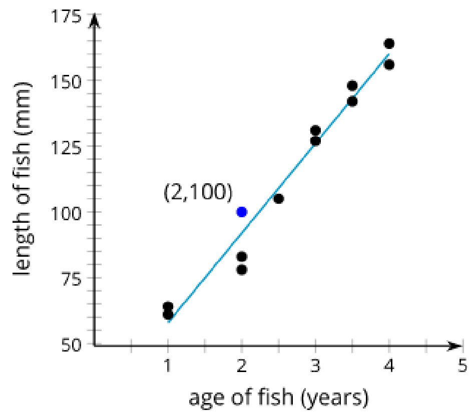
$$y = 91.94\ mm$$

To find the residual, we take the actual length of the two-year-old fish and subtract the length that our model predicts:

$$100 - 91.94 = 8.06\ mm$$

The residual of $8.06\ mm$ means the actual fish is about 8 millimeters longer than the linear model predicts for a fish of that same age.

$$y = 34.08x + 23.78$$



When the point on the scatter plot is above the line, it has a positive residual. When the point on the scatter plot is below the line, the residual is a negative value. A line that has smaller residuals would be more likely to produce predictions that are close to the actual value.

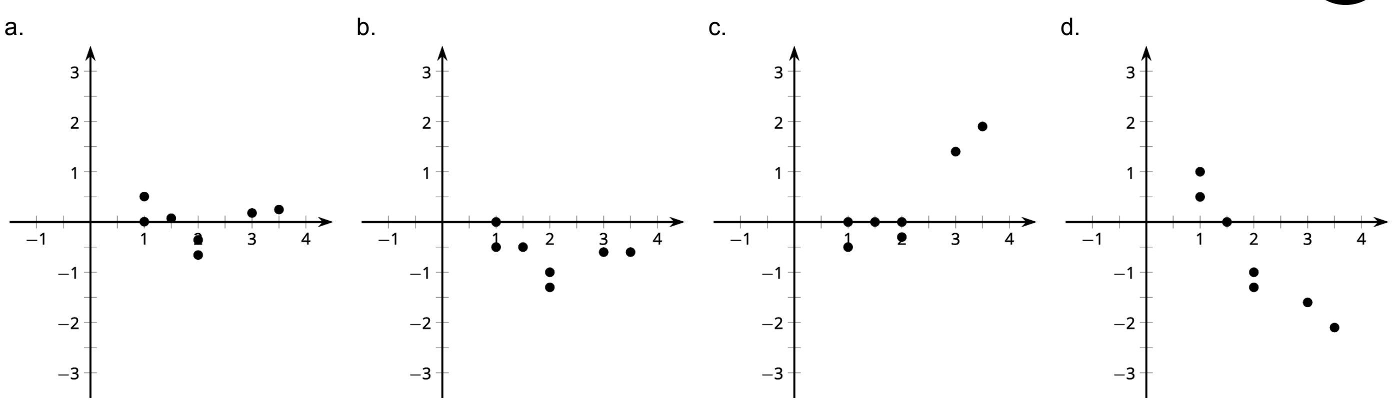
Cool-down: Deciding from Residuals (*5 minutes*)

Addressing: NC.M1.S-ID.6b; NC.M1.S-ID.8

Cool-down Guidance: Points to Emphasize
If students struggle here, select sample student work to discuss and assign practice problems to revisit this topic. Residuals will not be discussed again in this unit.

Cool-down

Each residual plot was made using a different line to fit the same set of data. Which graph is most likely to represent the residuals from the best fit line? Explain your reasoning.



Student Reflection: Consider your learning to identify a line of bad fit, good fit, best fit as well as residuals.

a. What tools best support your learning?

b. What tools were least supportive of your learning?

DO THE MATH

INDIVIDUAL STUDENT DATA

SUMMARY DATA

NEXT STEPS

TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

What opportunities are you giving students to reflect on their understanding of the mathematical content?

Practice Problems

1. Han creates a scatter plot that displays the relationship between the number of items sold, x, and the total revenue, y, in dollars. Han creates a line of best fit and finds that the residual for the point $(12, 1000)$ is 75. The point $(13, 930)$ has a residual of -40. Interpret the meaning of -40 in the context of the problem.

2. The line of best fit for a data set is $y = 1.1x + 3.4$. Find the residual for each of the coordinate pairs, (x, y).

   a. $(5, 8.8)$

   b. $(2.5, 5.95)$

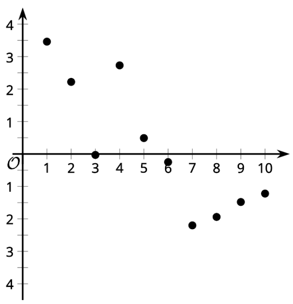
   c. $(0, 3.72)$

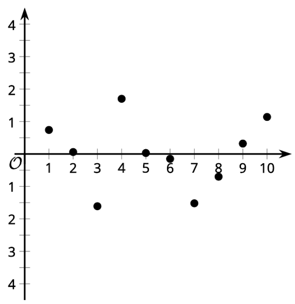
   d. $(1.5, 5.05)$

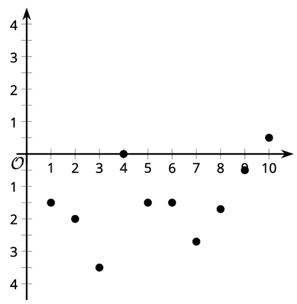
   e. $(-3, 0)$

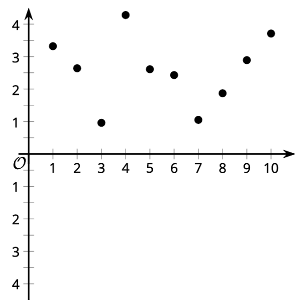
   f. $(-5, -4.86)$

3. Plots of the residuals for four different models of the same data set are displayed. Which of the following plots fit the data **best**?

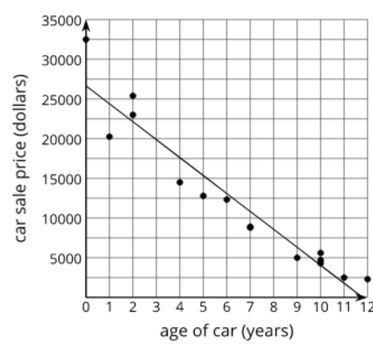
a.


b.


c.


d.


4. A local car salesperson created a scatter plot to display the relationship between a car's sale price in dollars, y, and the age of the car in years, x. The scatter plot and the line of best fit are displayed in the graph.



The car salesperson looks at the residuals for the car sales.

   a. For a car that is four years old, does the salesperson sell above or below her average selling price? Explain your reasoning.

   b. For a car that is 12 years old, does the salesperson sell above or below her average selling price? Explain your reasoning.

5. (*Technology required.*) Data about the outside temperature and gas used for heating a building are given in the table.

Use technology to create the line of best fit for the data. The data can be found on the Practice Problem tab: https://bit.ly/U4L7DataSet.

Temperature (deg F) x	Gas usage (therms) y
58	5,686
62	7,373
64	5,805
67	5,636
70	3,782
73	3,976
74	3,351
74	3,396
75	2,936
73	3,078
65	4,549
59	7,022
58	6,106
62	4,566
64	4,608
67	5,790
70	6,501
73	3,843

   a. What is the equation of the line of best fit for this data? Round numbers to the nearest whole number.

   b. What is the slope of the line of best fit? What does it mean in this situation?

   c. What does the line of best fit predict for gas usage when the outside temperature is 59 degrees Fahrenheit?

   d. How does the actual gas usage compare to the predicted gas usage when the outside temperature is 59 degrees Fahrenheit?

(From Unit 4, Lesson 4)

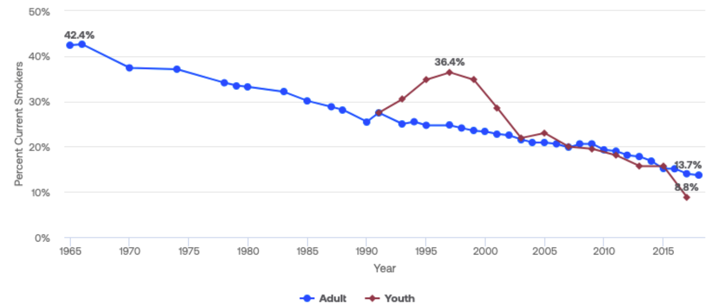
6. The graph below, from analysis of Center for Disease Control data by the American Lung Association, shows trends in cigarette smoking from 1965 to 2018.

   Based on the graph:

   a. Describe the pattern of cigarette smoking in adults and youth over the past 55 years.

   b. Would a linear model be a good fit for the cigarette smoking rates for adults? For youth? Explain your answer.

   (From Unit 4, Lesson 1)



(Image source[3])

7. One equation in a system of linear equations is $2x + 3y = -10$. If the system has no solutions, what could be the other equation?

   (From Unit 3)

8. Solve for x:

   a. $3.5x + 9 = 5x - 18$

   b. $3.5x + 9 > 5x - 18$

   c. For questions a and b, is 18 a solution to a, b, both, or neither? Why?

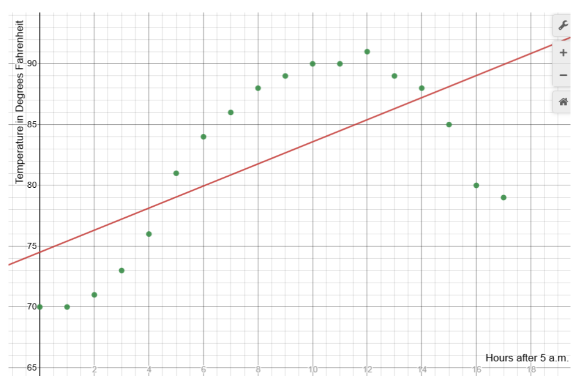
   d. For questions a and b, is 20 a solution to a, b, both, or neither? Why?

   (From Unit 2)

9. The scatterplot below represents the temperature every hour on June 21 in Charlotte, NC, starting at 5:00 a.m. Andre draws a line of best fit to try to predict the temperature the next morning.

   Is Andre's line a good model to predict the temperature on the morning of June 22? Explain your reasoning.

   (Addressing 8.SP.2)



[3] American Lung Association. *Overall Tobacco Trends*. https://www.lung.org/research/trends-in-lung-disease/tobacco-trends-brief/overall-tobacco-trends

Lesson 8: Causal Relationships

PREPARATION

Lesson Goal	Learning Target
• Investigate the relationship between two variables to analyze whether or not the relationship is causal.	• I can look for connections between two variables to analyze whether or not there is a causal relationship.

Lesson Narrative

The mathematical purpose of this lesson is to understand that the relationship between variables can be, but is not always, a causal relationship. **Association** indicates a relationship between two or more variables.[1] A **causal relationship** is one in which a change in one of the variables directly causes a change in the other variable. The work of this lesson connects to previous work because students interpreted the relationship between two variables using the correlation coefficient. The work of this lesson connects to the work of the upcoming modeling lesson because students will analyze bivariate data and draw conclusions from their analysis.

When students determine that there is a causal relationship, they are attending to precision (MP6), because they are refining their language to be more precise. When students analyze the relationship between two variables to determine whether they are causal or not, they are modeling with mathematics (MP4).

What connections might students make between this lesson and their own lives, personal interests, other subjects, and/or current events?

Focus and Coherence

Building On	Addressing
NC.M1.S-ID.6: Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.	**NC.M1.S-ID.8:** Analyze patterns and describe relationships between two variables in context. Using technology, determine the correlation coefficient of bivariate data and interpret it as a measure of the strength and direction of a linear relationship. Use a scatter plot, correlation coefficient, and a residual plot to determine the appropriateness of using a linear function to model a relationship between two variables. **NC.M1.S-ID.9:** Distinguish between association and causation.

[1] From NC Department of Public Instruction https://bit.ly/NCMath1InstructionalSupportsTool

Agenda, Materials, and Preparation

- **Bridge** (*Optional, 5 minutes*)
- **Warm-up** (*5 minutes*)
- **Activity 1** (*15 minutes*)
- **Activity 2** (*10 minutes*)
  - Find Your Cause card sort (print 1 copy per group of students and cut up in advance)
- **Activity 3** (*Optional, 20 minutes*)
- **Lesson Debrief** (*5 minutes*)
- **Cool-down** (*5 minutes*)
  - M1.U4.L8 Cool-down (print 1 copy per student)

LESSON

Bridge (*Optional*, *5 minutes*)

Instructional Routine: Notice and Wonder

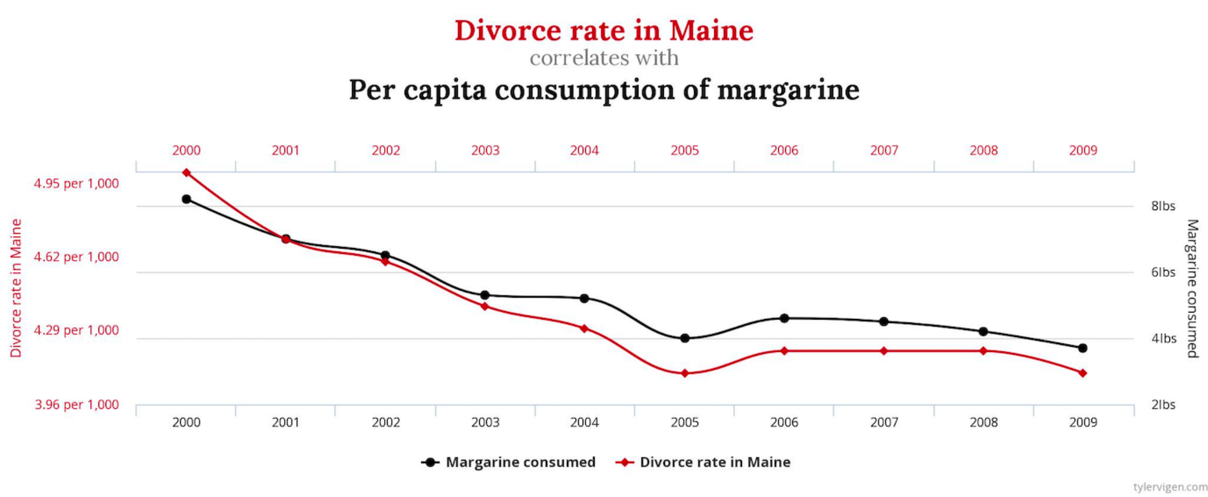
The purpose of this bridge is to allow students to articulate, in their own words, that one trend similar in pattern to another does not necessarily imply that the first causes the second or vice versa. The categories in this bridge were intentionally chosen because of their lack of obvious relationship. While students will be interacting with linear two-variable scatter plots in this lesson, building from NC.8.SP.1, it is also important to honor students' experience with various types of graphs outside of their studies in math classes. Students can leverage the arguments they make here about what a causal relationship means to aid their discussions in the lesson's activities.

Using the *Notice and Wonder* routine, ask students what they notice and wonder as they look over the task. As students share, listen for informal language they use to indicate an association or a causation. Leverage this informal language when formal definitions are presented later in this lesson.

Student Task Statement

The following graph displays the divorce rate in Maine and the number of pounds of margarine consumed per person each year. Study the graph.[2]

What do you notice? What do you wonder?



[2] Adapted from Tyler Vigen

DO THE MATH	PLANNING NOTES

Warm-up: Used Car Relationships (*5 minutes*)

Addressing: NC.M1.S-ID.8	**Building Towards:** NC.M1.S-ID.9

The mathematical purpose of this activity is for students to describe the relationship between variables using mathematical terminology, such as strong or weak relationship and positive or negative relationship. These terms were defined in Lesson 2. Students must reason abstractly and quantitatively (MP2) to determine the type of relationship.

Step 1

- Ask students to arrange themselves in pairs or use visibly random grouping.
- Remind students there are many possible answers for the questions.
- After 1–2 minutes of quiet work time, ask students to compare their responses to their partner's and decide if they are both correct, even if they are different. Follow with a whole-class discussion.
- To help students understand some of the context, explain that, for many modern cars, it is recommended that the oil be changed every 5,000 miles driven or every five months.

Student Task Statement

Describe the strength and direction of the relationship you expect for each pair of variables. Explain your reasoning.

1. Used car price and original sale price of the car
2. Used car price and number of cup holders in the car
3. Used car price and number of oil changes the car has had
4. Used car price and number of miles the car has been driven

Step 2

- The purpose of this discussion is for students to discuss the strength of relationships in preparation for having students distinguish between causal and non-causal relationships. Encourage the use of the terms "strong" or "weak" relationship and "positive" or "negative" relationship in the discussion.
- Focus the discussion on the reasoning for the questions 3 and 4 and how they are similar and different. Here are some questions for discussion.
  - "How is the relationship between the car price and number of oil changes similar to the relationship between the car price and number of miles driven?" (They both have a strong, negative relationship.)

– "How is the relationship between the car price and number of oil changes different from the relationship between the car price and number of miles driven?" (The number of miles the car has been driven seems more directly related to the price than the number of oil changes. It probably has a stronger negative relationship than the number of oil changes.)

DO THE MATH	PLANNING NOTES

Activity 1: Causation or Association? (*15 minutes*)

Instructional Routine: Aspects of Mathematical Modeling	
Building On: NC.M1.S-ID.6	**Addressing:** NC.M1.S-ID.9

The purpose of this activity is for students to describe how two variables are related, and to determine whether or not there is a causal relationship. Students should begin to recognize that some variables may be related, but one does not cause the other to change. At this point, the mathematics of scatter plot analysis cannot decide whether there is a causal relationship. The relationship must be thought through carefully to decide, based on the situation, whether the related variables have a causal relationship. As students analyze the relationships, they engage with *Aspects of Mathematical Modeling* (MP4).

RESPONSIVE STRATEGIES

Create a display of important terms and vocabulary. After the first question, invite students to suggest language or diagrams to include that will support their understanding of causal relationships. Examples to include can be phrases, such as, "increased rain causes people to wear jackets." The two variables can be color-coded in two different colors to reveal the relationship. Both the axes and the labels can be highlighted with the color that connects it to the type of variable it addresses.

As students compare their responses with their partner, display two sentence frames: "I notice we both think___" and "We described ___ differently because ____."

Supports accessibility for: Conceptual processing; Language

Step 1

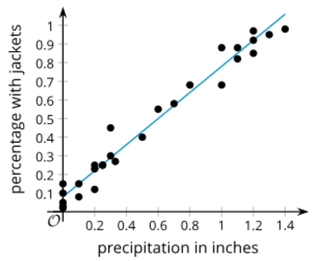
- Have students remain with the same partner.
- Tell students there are many possible answers for the questions.
- After 5 minutes of quiet work time, ask students to compare their responses to their partner's and decide if they are both correct, even if they are different.

Monitoring Tip: As students compare their responses, monitor for phrases students use that indicate causal relationships (for example: an increase in _________ causes a decrease in ________). Prepare to have these students share with the class in the whole-group discussion following the activity.

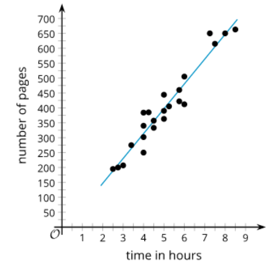
Student Task Statement

Each of the scatter plots shows a strong relationship. Write a sentence or two describing how you think the variables are related.

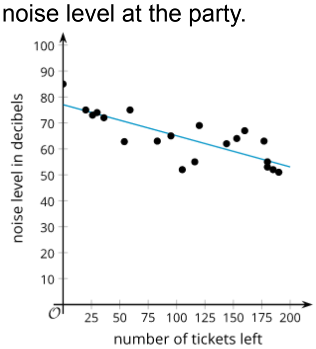
1. During the month of April, Elena keeps track of the number of inches of rain recorded for the day and the percentage of people who come to school with rain jackets.



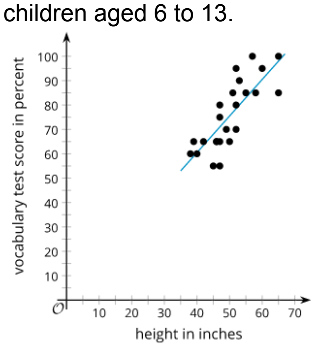
2. A school book club has a list of 100 books for its members to read. They keep track of the number of pages in the books the members read from the list and the amount of time it took to read the book.



3. Number of tickets left for holiday parties at a venue and noise level at the party.



4. The height and score on a test of vocabulary for several children aged 6 to 13.



Step 2

- Select several students to share their reasoning for the relationship between the variables.
- For each pair of variables, ask students what might have *caused* the variables to be linked. In some cases (questions 1 and 2), one of the variables causes a change in the other. In other cases (questions 3 and 4), an additional variable or situation is the cause of the change. Here are some questions for discussion:
    - “Why does an increase in precipitation cause an increase in the percentage of people wearing rain jackets?” (When it rains, people usually wear coats to stay dry.)
    - “Does the time it takes to read a book cause the number of pages in the book to increase?” (No.) “Does an increase in the number of pages in a book cause the time it takes to read a book to increase? ” (Yes, if you switch the axes on the graph, you can see that relationship. It is causal because it takes longer to read more pages.)
    - “The relationship in the height and test score graphs appears to be strong. Does an increase in height cause an increase in test scores?” (It is not the height that causes the increase in test score, but probably the age or grade level of the students might be causing the increase. Age and grade level are positively related to height, so that likely explains the relationship seen between the two variables.)
- Tell students that, since most people are used to seeing independent variables on the x-axis and dependent variables on the y-axis, the convention is to put the causal variable on the x-axis and the other variable on the y-axis when one of the variables does cause the other to change. The book club scatter plot should probably have the axes switched to meet the convention.

DO THE MATH	PLANNING NOTES

Activity 2: Find Your Cause (*10 minutes*)

Instructional Routines: Card Sort; Stronger and Clearer Each Time (MLR1)
Addressing: NC.M1.S-ID.9

This *Card Sort* activity allows students to practice distinguishing between causal relationships and relationships that have association but are not causal.

Review the Find Your Cause card sort for planning purposes.

Step 1

- Ask students to remain in pairs from the previous activity.
- Tell students that "people often use the phrase 'correlation, not causation' or 'association, not causation' to refer to these situations in which there is a relationship, but it is not a causal relationship. A causal relationship means that a change in one of the variables actually *causes* a change in the other variable."
- Give students several minutes to categorize their cards individually and then write some initial notes about why they believe each card belongs in the selected category.

RESPONSIVE STRATEGIES

Encourage and support opportunities for peer collaboration. Prior to directing student attention to the task, invite students to brainstorm various pairs of variables (both with causation and without). Encourage students to free-associate and build off of each other's ideas. When students share their work with a partner, display sentence frames to support conversation such as: "____ reminds me of ____ because . . ." and "____'s idea reminds me of . . ." After students have developed a list of several options, share the conditions list and then encourage them to analyze and select which one best fits each condition.

Supports accessibility for: Language; Social-emotional skills

Step 2

Use the *Stronger and Clearer Each Time* routine to help students clarify their ideas through conversation.

- Students should first check to see if they agree with each other's sorting of cards. Then partners should take turns explaining the reasoning for one of their choices, and providing each other with feedback on their explanations.
- Provide listeners with prompts for feedback that will help their partner add detail to strengthen and clarify their ideas. For example, students can ask their partner: "Why do you think these are not causal?"
- Have students repeat the process with another partner.
- Finally, provide students with 2–3 minutes to revise their initial notes based on feedback and ideas from the conversations with their peers. This will help students understand and communicate about relationships that may or may not be causal.

***Advancing Student Thinking*:** Students may still wonder how two variables can be associated without having a causal relationship. It may help to provide an example, such as sales of ice cream and sales of sunburn remedies. Ask students why these variables might be related and whether increasing one would cause the other variable to increase. Ask students to think of something that might cause two distinct outcomes to result.

Student Task Statement

1. Your teacher will give you a set of cards. Individually, categorize each relationship as causation or association without causation. Once finished, share your categorization with your partner.

   For each category, explain to your partner how you know each card belongs in that category. Carefully listen to your partner as they explain their selections. If you disagree, discuss your thinking and work to reach an agreement. After coming to an agreement, record the examples in the agreed-upon category.

Causation	Association without causation

2. If time allows, brainstorm with your partner an example of a relationship that has association but not causation.

Are You Ready For More?

1. Look through news articles or advertisements for claims of causation or association. Find two or three claims and read or watch the articles or the advertisement. Answer these questions for each of the claims.
   a. What is the claim?
   b. What evidence is provided for the claim?
   c. Does there appear to be evidence for causation or correlation? Explain your thinking.
2. Choose the claim with the least or no evidence. Describe an experiment or other way that you could collect data to show correlation or causation.

DO THE MATH

PLANNING NOTES

Activity 3: Fossil Puzzle (*Optional*, *20 minutes*)

Addressing: NC.M1.S-ID.6; NC.M1.S-ID.7; NC.M1.S-ID.8; NC.M1.S-ID.9

The mathematical purpose of this activity is for students to summarize, interpret, and draw conclusions from bivariate data using scatter plots, best fit lines, residuals, and correlation coefficients. Students create a linear model based on the approximate lengths of the humerus bones and heights given in a table. The model is then used to approximate the height of an ancient human based on the length of a found humerus bone.

Making graphing technology available gives students an opportunity to choose appropriate tools strategically (MP5). By collecting their own data and using a best fit line to find additional information, students are modeling with mathematics (MP4).

Step 1

- Ask students to arrange themselves in pairs or use visibly random grouping.
- Present the task to students and provide 2 minutes of quiet think time for students to brainstorm different ways that they could answer the question.
- Ask students to share their ideas with the class. The remaining time should be used by students to analyze, summarize, and interpret the data.

Student Task Statement

An anthropologist finds a fossilized humerus bone of an ancient human ancestor. The humerus is an arm bone running from the shoulder to the elbow. It is 24 centimeters in length.

a. How tall do you think this ancient human was?

b. In order to estimate the height of this ancient human, Kiran decided to measure the humerus of some of his friends and family and see if that would give him an idea.

The table below shows the length of the humerus (in centimeters) and the height (in inches) of 20 of Kiran's friends and family. This data can be found on the Activity 3 tab: https://bit.ly/U4L8DataSets.

Humerus (cm)	33	37	35.5	31.5	30.3	30.5	30.5	28	35	36	28	29	30.5	25.4	33.3	23	26	33	27.8	28.4
Height (in)	64	73	70	67	61.8	67	66	63	66	72	59	62	73	64	68	65	67	75	56	60

c. Use the data Kiran collected to estimate the height of the ancient human.

d. How confident are you in your answer? What information helped you determine your confidence?

e. How did you use mathematics to estimate the height of the ancient human?

Step 2

- Facilitate a whole-class discussion focused on how students used mathematics to justify their estimation. Ask students:
  - "How confident are you in your answer? What information helped you determine your confidence?" (Not very confident. Since the correlation coefficient is near 0.6, there is only a moderate relationship between height and humerus length. Additionally, this ancient human ancestor may have a different anatomy—for example, apes tend to have proportionally longer arms than humans do.)
  - "How did you use mathematics to estimate the height of the ancient human?" (I made a scatter plot to determine whether or not a linear model was appropriate, and then computed a line of best fit. I substituted 24 centimeters into my line of best fit and obtained my answer.)

DO THE MATH	PLANNING NOTES

Lesson Debrief (*5 minutes*)

The purpose of this lesson was for students to gain a deeper understanding of what it means for variables to have a causal relationship.

Choose what questions to focus the discussion on, whether students should first have an opportunity to reflect in their workbooks or talk through these with a partner, and what questions will be prioritized in the full class discussion. Discuss questions such as:

- "How can you determine if there is a causal relationship between two variables? Explain your reasoning." (To determine a causal relationship, you need to think about the context and determine if a change in one variable causes the other variable to change.)
- "Mai states that the relationship between the number of miles driven in a taxi and the price of the taxi ride is a causal relationship. Do you agree with Mai? (I agree with her. It makes sense that the farther you go in a taxi, the more you will be charged.)
- "Jada states that the relationship between the size of a pasture and the number of cows kept at various farms is a causal relationship. Do you agree with Jada? Explain your reasoning." (I do not agree. The increase in pasture size does not cause an increase in the number of cows [nor does an increase in number of cows make the pasture larger]. The increase in land might mean there is more food for the cows, but it is the farmer who decides how many cows there are on a farm. OR: I agree. Farmers with bigger pastures have room to acquire more cows, whereas farmers with smaller pastures cannot.)

PLANNING NOTES

Student Lesson Summary and Glossary

Humans are wired to look for connections and then use those connections to learn about the world around them. One way to notice connections is by looking for a pair of variables with a relationship. For example, if we notice that people who tend to eat many calories also have a higher chance of having a heart attack, we might wonder if lowering our calorie intake would improve our health.

One common mistake people tend to make using statistics is to think that all relationships between variables are causal. Scatter plots can only show a relationship between the two variables. To determine if change in one of the variables actually causes a change in the other variable, or has a causal relationship, the context must be better understood and other options ruled out.

For example, we might expect to see a strong, positive relationship between the number of snowboard rentals and sales of hot chocolate during the months of September through January. This does not mean that an increase in snowboard rentals causes people to purchase more hot chocolate. Nor does it mean that increased sales of hot chocolate cause people to rent snowboards more. More likely there is a third variable, such as colder weather, that might be causing both variables to increase at the same time. In this instance, there is an **association** between snowboard rentals and hot chocolate sales but not necessarily a causation.

Association: A relationship between two or more variables.

On the other hand, sometimes there is a **causal relationship**. A strong, positive relationship between hot chocolate sales and small marshmallow sales may be linked because people buying hot chocolate may want to add small marshmallows to the drink, so an increase in the sales of hot chocolate actually causes the marshmallow sale increase.

Causal relationship: A relationship in which a change in one of the variables causes a change in the other variable.

Cool-down: Just Cause (*5 minutes*)

Addressing: NC.M1.S-ID.9

Cool-down Guidance: Press Pause
If students struggle to describe relationships between variables, spend the first 5 minutes of Lessons 9 & 10 highlighting key ideas from cool-downs in Lessons 7–8.

Cool-down

For each pair of variables, decide whether there is:

- a very weak or no relationship
- a strong relationship that is not a causal relationship
- a causal relationship

Pair of variables	Explain your reasoning
1. Number of snow plows owned by a city and mitten sales in the city	
2. Number of text messages sent per day by a person and number of shirts owned by the person	
3. Price of a pizza and number of calories in the pizza	
4. Amount of gas used on a trip and number of miles driven on the trip	

Student Reflection: Consider your own experiences. What are one or two examples of associations you have seen or experienced? What are one or two examples of causal relationships you have seen or experienced?

DO THE MATH

INDIVIDUAL STUDENT DATA

SUMMARY DATA

NEXT STEPS

TEACHER REFLECTION

What went well in the lesson? What would you do differently next time? What happened today that will influence the planning of future lessons?

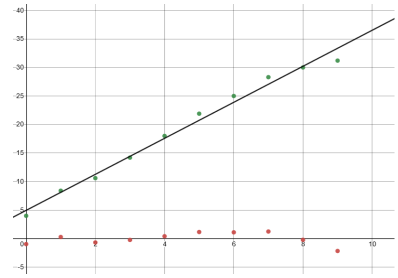
As students worked together today, where did you see evidence of the mathematical community established over the course of the school year?

Practice Problems

1. Priya creates a scatter plot showing the relationship between the number of steps she takes and her heart rate. The correlation coefficient is 0.88.
    a. Describe the association between Priya's steps and her heart rate.
    b. Do either of the variables cause the other to change? Explain your reasoning.
2. Kiran creates a scatter plot showing the relationship between the number of students attending drama club and the number of students attending poetry club each week. The correlation coefficient is -0.36.
    a. Describe the association between the variables.
    b. Do either of the variables cause the other to change? Explain your reasoning.
3. A news website shows a scatter plot with a negative relationship between the amount of sugar eaten and happiness levels. The headline reads, "Eating sugar causes happiness to decrease!"
    a. What is wrong with this claim?
    b. What is a better headline for this information?

4. The graph below shows a scatter plot, the line of best fit, and the residual plot. Add a point to the scatter plot at $x = 10$, and plot the corresponding residual. How did you decide where to plot your point and the residual?

   (From Unit 4, Lesson 7)



5. (*Technology required.*) The following table shows the average yearly price for one gallon of gas from the years 2000–2010, according to the U.S. Energy Information Administration. Data can be found on the right and on the practice problem tab: https://bit.ly/U4L8DataSets.

Year	Price per gallon
2000	$1.52
2001	$1.46
2002	$1.39
2003	$1.60
2004	$1.90
2005	$2.31
2006	$2.62
2007	$2.84
2008	$3.30
2009	$2.41
2010	$2.84

   a. Determine the line of best fit and correlation coefficient for the gas prices. Assume the year 2000 is where $x = 0$.

   b. Explain the meaning, in context, of the slope and vertical intercept of the line of best fit.

   c. Is the line of best fit a good model for the data? Explain your reasoning.

   (From Unit 4, Lessons 2–4)

6. The number of miles driven, x, and the number of gallons remaining in the gas tank, y, have a strong negative relationship.

   Explain what it means to have a **strong negative** relationship in this context.

   (From Unit 4, Lesson 2)

7. (*Technology required.*) Data can be found to the right and on the practice problem tab: https://bit.ly/U4L8DataSets.

x	y
10.2	31
10.4	27
10.5	29
10.5	30
10.5	31
10.6	26
10.8	25
10.8	26
10.9	27
11	24
11.2	22

   a. What is an equation of the line of best fit?

   b. What is the value of the correlation coefficient?

   (From Unit 4, Lesson 2)

8. Here are the heights in inches of the players on the boys basketball team at Fanwood High School:

   73, 71, 74, 83, 72, 72, 68, 71

   a. Calculate the five-number summary for the heights of the players on the team.

   b. Which height is a potential outlier? Describe what that means in relation to the other players.

   c. How would the standard deviation be affected if the outlier was removed? How do you know?

   (From Unit 1)

Lessons 9 & 10: Mathematical Modeling[1]

PREPARATION

Lesson Goals	Learning Targets
• Use mathematical tools to represent, interpret, analyze, generalize, communicate about, and make predictions about how real-world quantities vary in relation to each other. • Use observational and experimental data to test, improve, and validate mathematical models.	• I can use mathematics to model real-world situations. • I can test and improve mathematical models for accuracy in representing and predicting real things.

Lesson Narrative

Modeling is the link between the mathematics students learn in school and the problems they will face in college, career, and life. Time spent on modeling is crucial, as it prepares students to use math to handle technical subjects in their further studies, and problem solve and make decisions that adults regularly encounter in their lives.

For Lessons 9 and 10, there are four choices of modeling prompts: Modeling Prompts #1 and #2 are available in Unit 2, and Modeling Prompts #3 and #4 are provided here. If Modeling Prompt #1 was not done in Unit 2, it is recommended to begin there as it is a good introduction to the expectations around modeling.

Remind students what modeling is and what is expected of them as a modeler using the following resources and guidance:

Organizing Principles about Mathematical Modeling

- The purpose of mathematical modeling in school mathematics courses is for students to understand that they can use math to better understand things in the world that interest them.
- Mathematical modeling is different from solving word problems. It often feels like initially there is not enough information to answer the question. There should be room to interpret the problem. There ought to be a range of acceptable assumptions and answers. Modeling requires genuine choices to be made by the modeler.
- It is expected that students have support from their teacher and classmates while modeling with mathematics. It is not a solitary activity. Assessment should focus on feedback that helps students improve their modeling skills.

Things the Modeler Does When Modeling with Mathematics (NGA 2010)

1. *Pose a problem* that can be explored with quantitative methods. Identify variables in the situation and select those that represent essential features.
2. *Formulate a model*: Create and select geometric, graphical, tabular, algebraic, or statistical representations that describe relationships between variables.
3. *Compute*: Analyze these relationships and perform computations to draw conclusions.
4. *Interpret* the conclusions in terms of the original situation.
5. *Validate* the conclusions by comparing them with the situation. Iterate if necessary to improve the model.
6. *Report* the conclusions and the reasoning behind them.

It's important to recognize that in practice, these actions don't often happen in a nice, neat order.

[1] Adapted from IM 9–12 Math Algebra 1 Modeling Prompts https://curriculum.illustrativemathematics.org/HS/index.html, copyright 2019 by Illustrative Mathematics. Licensed under the Creative Commons Attribution 4.0 license https://creativecommons.org/licenses/by/4.0/.

Preparing for a Modeling Prompt

Ideas for Setting Up an Environment Conducive to Modeling

- Provide plenty of blank whiteboard or chalkboard space for groups to work together comfortably. “Vertical non-permanent surfaces” are most conducive to productive collaborative work. “Vertical” means on a vertical wall, which is better than horizontally on a tabletop, and “non-permanent” means something like a dry erase board, which is better than something like chart paper (Liljedahl 2016).
- Ensure that students have easy access to any tools that might be useful for the task. These might include:
  - supply table containing geometry tools, calculators, scratch paper, graph paper, dry erase markers (ideally a different color for each group member), post-its
  - electronic devices to access digital tools (like graphing technology, dynamic geometry software, or statistical technology)
- Think about how to help students manage the time that is available to work on the task. For example:
  - Display a countdown timer for intermittent points in the lesson when groups are asked to summarize their progress.
  - Decide what time to ask groups to transition to writing down their findings in a somewhat organized way (perhaps 15 minutes before the end of the class).

How to Prepare and Conduct the Modeling Lesson

- Decide which version of the prompt students will receive, based on the lift-analysis, timing, and access to data.
- Have data ready to share if planning to give it when students ask.
- Decide if students will be offered a template for organizing modeling work.
- Decide to what extent students are expected to iterate and refine their model. The amount of time available can influence how much time students have to refine their model. If time is short, students may not engage as much in that part of the modeling cycle. WIth more time, it is more reasonable to expect students to iterate and refine their model once or even several times.
- Decide how students will report their results. Again, if time is short, this may be a rough visual display on a whiteboard. If more time is available, students might create a more formal report, slideshow, blog post, poster, mockup of an artifact like a letter to a specific audience, smartphone app, menu, or set of policies for a government entity to consider. One way to scaffold this work is to ask students to turn in a certain number of presentation slides: one that states the assumptions made, one that describes the model, and one or more slides with their conclusions or recommendations.
- Develop task-specific “look-fors” for each dimension of the provided rubric. What do you anticipate and hope to see in student work?

Organizing Students Into Teams or Groups

- Mathematical modeling is not a solitary activity. It works best when students have support from each other and their teacher.
- Working with a team can make it possible to complete the work in a finite amount of class time. For example, the team may decide it wants to vary one element of the prompt and compute the output for each variation. What would be many tedious calculations for one person could be only a few calculations for each team member.
- The members of good modeling groups bring a diverse set of skills and points of view. Create and share a Multiple Abilities List with students.
- Scramble the members of modeling teams often, so that students have opportunities to play different roles.

Ways to Support Students While They Work on a Modeling Prompt

- Coach students on ways to organize their work.
- Provide a template to help students organize their thinking. Over time, some groups may transition away from needing to use a template.
- Engage students in the *Three Reads* instructional routine to ensure comprehension of the prompt.
- Remind students of the variety of tools that are available to them.
- If students get stuck or run out of ideas, help move them forward with a question that prompts them to focus on a specific part of the modeling cycle. For example:
  - “What quantities are important? Which ones change and which ones stay the same?”
  - “What information do you know? What information would it be nice to know? How could you get that information? What reasonable assumption could you make?”
  - “What pictures, diagrams, graphs, or equations might help people understand the relationships between the quantities?”
  - “How are you describing the situation mathematically? Where does your solution come from?”
  - “Under what conditions does your model work? When might it not work?”
  - “How could you make your model better? How could you make your model more useful under more conditions?”
  - “What parts of your solution might be confusing to someone reading it? How could you make it more clear?”

How to Interpret the Provided Lift Analysis of a Modeling Prompt

For most mathematical modeling prompts, different versions are provided. Each version is analyzed along five impactful dimensions that vary the demands on the modeler (OECD 2013). Each of the attributes of a modeling problem is scored on a scale from 0–2. A lower score indicates a prompt with a “lighter lift” for students and teachers: students are engaging in less open, less authentic mathematical modeling. A higher score indicates a prompt with a “heavier lift” for students and teachers: students are engaging in more open, more authentic mathematical modeling. This matrix shows the attributes that are part of our analysis of each mathematical modeling prompt. Though not all the attributes have the same impact on what teachers and students do, for the sake of simplicity, they are all weighted the same when they are averaged.

Attribute		DQ Defining the Question	QI Quantities of Interest	SD Source of Data	AD Amount of Data given	M The Model
Lift	Light lift (0)	Well-posed question	Key variables declared	Data are provided	Modeler is given all the information they need and no more	Model is given in the form of a mathematical representation
	Medium lift (1)	Elements of ambiguity; prompt might suggest ways assumptions could be made	Key variables suggested	Modelers are told what measurements to take or data to look up	Some extra information is given and modeler must decide what is important; or, not enough information is given and modeler must ask for it before teacher provides it	Modeler must sift through lots of given information and decide what is important; or, not enough information is given and modeler must make assumptions, look it up, or take measurements
	Heavy lift (2)	Freedom to specify and simplify the prompt; modeler must state assumptions	Key variables not evident	Modelers must decide what measurements to take or data to look up	Modeler must sift through lots of given information and decide what is important; or, not enough information is given and modeler must make assumptions, look it up, or take measurements	Careful thought about quantities and relationships or additional work (like constructing a scatter plot or drawing geometric examples) is required to identify type of model to use

Each version of a mathematical modeling prompt is accompanied by an analysis chart that looks like this sample:

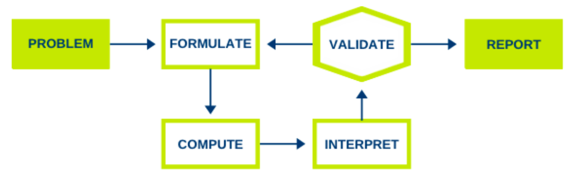
Attribute	DQ	QI	SD	AD	M	Avg
Lift	0	1	0	0	2	0.6

There are other features of a mathematical modeling prompt that could be varied. In the interest of not making things too complex, there are only five dimensions included in the lift analysis. However, a prompt could be additionally modified on one of these dimensions:

- whether the scenario is posed with words, a highly-structured image or video, or real-world artifacts like articles or authentic diagrams
- presenting example for student to explore before they are expected to engage with the prompt, versus the prompt suggesting that the modeler generate examples or expecting the modeler to generate examples on their own
- whether the prompt makes decisions about units of measure or expects the modeler to reconcile units of measure or employ dimensional thinking
- whether a pre-made digital or analog tool is provided, instructions given for using a particular tool, use of a particular tool is suggested, or modelers simply have access to familiar tools but are not prompted to use them
- whether a mathematical representation is given, suggested, or modelers have the freedom to select and create representations of their own choosing

Advice on Modeling

These are some steps that successful modelers often take and questions that they ask themselves. You don't necessarily have to do all of these steps, or do them in order. Only do the parts that you think will help you make progress.



Understand the Question
Think about what the question means before you start making a strategy to answer it. Are there words you want to look up? Does the scenario make sense? Is there anything you want to get clearer on before you start? Ask your classmates or teacher if you need to.

Refine the Question
If necessary, rewrite the question you are trying to answer so that it is more specific.

Estimate a Reasonable Answer
If you don't have enough information to decide what's reasonable, try to come up with an answer that would be too low, and an answer that would be too high.

Identify Unknowns
- What are the meaningful quantities in this situation? Write them down.
- What information would be useful to know? In order to get that information, you could: look it up, take a measurement, or make an assumption.

Gather Information
Write down any of the unknown information that you find. As you work, organize your information in a way that makes sense to you.

Experiment!
Try different ideas to make progress toward answering your question. If you are stuck, think about:
- Helpful ways to organize the information you have or organize your work
- Questions you *can* answer using the information you have
- Ways to represent mathematical relationships or sets of data (tables, equations, scatter plots, graphs, statistical plots)
- Tools that are available for representing mathematics, both digital and analog

Check Your Reasoning
Do you have a first answer to your question? Great! See if it's reasonable.
- Make sure you can explain what the answer means in terms of the original problem.
- Check your precision: Is your answer overly precise (do you really need all those decimal places)? Not precise enough (were you overly aggressive with your rounding)?

Use and Improve Your Model
- Did you make assumptions or measurements? How can you express your model more generally, so that it would work for a range of numbers instead of the specific numbers you used?
- What are the limitations of your model? That is, what are some ways it is not realistic? Does it only work for certain inputs but not others? Are there any meaningful inputs affecting the outcome that are not accounted for? If possible, improve your model to take these into account.
- What are the implications of your model? That is, what should people or organizations do differently or smarter as a result of what your model shows? What would be effective ways to communicate with them?
- What are the areas for further research? That is, what new things are you wondering about that could be investigated, by you or someone else?

Modeling Rubric

Skill	Score			Notes or Comments
	Proficient	Developing	Needs Revisiting	
1. Decide What to Model	• Assumptions made are clearly identified and justified. Resulting limitations are stated when appropriate. • Variables of interest are clearly identified and chosen wisely, and appropriate units of measure are used.	• Assumptions are noted but lacking in justification or difficult to find. • Variables of interest are noted, but may lack justification, be difficult to find, or not be measured with appropriate units.	• No assumptions are stated. • No variables are defined.	
	To improve at this skill, you could: • Ask questions about the situation to understand it better • Check the assumptions you're making to see if they're reasonable (Try asking a friend, or imagining that you're a person involved in the scenario. Would those assumptions make sense to you?) • Double-check the variables you've identified: Are there other quantities in the situation that could vary? Is there something you've identified as a variable that is actually fixed or determined? (Remember that more abstract things like time and speed are also quantities.)			
2. Formulate a Mathematical Model	• An appropriate model is chosen and represented clearly. • Diagrams, graphs, etc. are clear and appropriately labeled.	Parts of the model are unclear, incomplete, or contain mistakes.	No model is presented, or the presentation contains significant errors.	
	To improve at this skill, you could: • Check your model more carefully to make sure it really fits well • Consider a wider variety of possible models, to find one that fits the situation better • Think about the situation more deeply before trying to find a model • Convince a skeptic: Pretend that you think your model is inadequate, or ask a friend to pretend to be skeptical of it. What would a skeptic find wrong with your model? Try to fix those things, or explain why they're not actually problems.			

Skill	Score			Notes or Comments
	Proficient	Developing	Needs Revisiting	
3. Use Your Model to Reach a Conclusion	• Solution is relevant to the original problem. • Reader can easily understand the reasoning leading to the solution. • Relevant details are included like units of measure.	Solution is not well-aligned to the original problem, or aspects of the solution are difficult to understand or incomplete.	No solution is provided.	
	To improve at this skill, you could: • Double-check your calculations: Show them to someone else to see if they agree, or take a break and look at your calculations again later • Make sure your calculations are justified by your model: Ask yourself how you decided what to calculate, and see if your reasoning matches up with your model • Think more deeply about what your conclusions mean in the original scenario: Imagine you're a person involved in the scenario, or explain your conclusions to someone else and see if they have questions			
4. Refine and Share Your Model	• The model's implications are clearly stated. • The limitations of the model and solution are addressed.	The limitations of the model and solution are addressed but lacking in depth or ignoring key components.	No interpretation of model and solution is provided.	
	To improve at this skill, you could: • Think more creatively about what your conclusions mean: Ask yourself "If I was involved in this situation, what would I understand better because of these conclusions? What would I want to do next?" • Be skeptical of your model: What don't you like about it, and what can you do to fix those things? • Explain your model to someone else: Tell them how it works and why it's good. If you're not sure how it works or why it's good, you might need to change it.			

What resources around mathematical modeling will you revisit with students? What do you hope they take away from them?

Agenda, Materials, and Preparation (Choose from Modeling Prompts #1–4)

- **Modeling Prompts #1 & #2: See Math 1 Unit 2 materials**
- **Modeling Prompt #3: A New Heating System**
  - Modeling Prompt #3 (print 1 copy per student)
  - Modeling Rubric (print 1 copy per student)
- **Modeling Prompt #4: College Characteristics**
  - Modeling Prompt #4 (print 1 copy per student)
  - Modeling Rubric (print 1 copy per student)
  - College Data for prompt 4A: https://bit.ly/CollegeData4A
  - College Data for prompt 4B: https://bit.ly/CollegeData4B

LESSON

Modeling Prompt #3: A New Heating System

In this modeling prompt, students will work on determining a new heating system for a homeowner that is most cost effective and efficient. There are two versions of this prompt: 3A and 3B. In 3A, students are required to research alternatives to the homeowner's heating system. In 3B, students are given alternative options to choose from. Determine in advance which Modeling Prompt (3A or 3B) students will receive, based on the lift-analysis, timing, and access to data. Whenever possible, students should engage with the prompt with the higher lift analysis.

Student Task Statement 3A Lift Analysis

Attribute	DQ	QI	SD	AD	M	Avg
Lift	2	1	2	2	1	1.6

Student Task Statement 3B Lift Analysis

Attribute	DQ	QI	SD	AD	M	Avg
Lift	1	0	0	0	1	0.4

Step 1 (*Optional;* review modeling materials as necessary)

- Display and have students reference the Advice on Modeling handout in their Student Workbook. Pass out the Modeling Rubric.
- Facilitate a discussion around modeling. Share some of the following ideas:
  - Modeling prompts are often expressed in words, but unlike word problems, modeling prompts challenge the modeler to make reasonable assumptions, decide what information is important, ask or research for more information if needed, think creatively within constraints, and consider the implications of the model.
  - The process of modeling is cyclical, and it does not end by producing a "correct answer."
  - A mathematical model expresses a simplified relationship in the real world; models can be rough but still useful.
  - Models can often be refined to represent the real-world relationship more accurately.
  - Responses to modeling prompts can vary widely, but they often contain certain pieces: assumptions, calculations, a mathematical model (stated with an equation or equations, with a graph, with a geometric diagram, or in words), conclusions, and generalizations.
- Provide time for students to ask clarifying questions. Students should have a good-enough understanding of the rubric; deep understanding of the rubric is not needed at this time.

Step 2

- Solicit information that students already know about heating buildings, which they will use in this task. Here are some possible questions for discussion:
  - "How are buildings heated?" (fireplaces, boilers, gas furnaces)
  - "What kinds of energy can be used to heat buildings?" (electricity, natural gas, solar power)
  - "When in the year do you think our school uses the most power for heating?" (probably right before or after winter break, when it's pretty cold and there are still people in the building for most of the day)
- Tell students that in this task, they will investigate energy costs for some different heating systems. They will begin by considering a homeowner's current heating system and compare some other systems the homeowner could use instead. Tell students that energy is measured in kilowatt-hours, which is abbreviated kWh.
- Give students a preview of some of the calculations they'll need to do by showing why the homeowner currently pays $975/year to heat the house. Tell students that the current system is 60% efficient, which means that for every 100 kWh it uses, it only produces 60 kWh of heat. Display this statement for all to see:

For every 100 kWh of energy a certain heating system uses, it produces 60 kWh of heat. If the system has to produce 11,700 kWh of heat to heat a house for the winter, how many kWh of energy will it use?

- Ask students to think about how they would find an answer to this question. They do not need to calculate an answer, only think of a strategy. After some quiet think time, ask students to share their thoughts with a partner.
  - Invite students to share strategies with the class. Use one of the suggested strategies to calculate the answer, and write the steps for all to see. Here is one possible way:

$$\frac{100\ kWh\ \text{input}}{60\ kWh\ \text{output}} = \frac{x\ kWh\ \text{input}}{11{,}700\ kWh\ \text{output}}$$

$$\frac{100\ kWh\ \text{input}}{60\ kWh\ \text{output}} \cdot 11,700\ kWh\ \text{output} = x\ kWh\ \text{input}$$

$$19,500\ kWh = x$$

- To find out how much this will cost the homeowner, we need to know how expensive the fuel is. If we assume natural gas costs $0.05 per kWh, then multiplying 19,500 by 0.05 will give us the total cost of $975.

Step 3

- Pass out the pre-determined appropriate blackline master Modeling Prompt #3 (3A or 3B) and Modeling Rubric if it has not already been distributed.
  - **Modeling Prompt 3A**: Students will need to do a lot of independent research for this task. One way to make the research less time-consuming would be to have students suggest other ways of heating the house, like solar panels or geothermal heating and cooling systems, and have each group investigate one of the ideas. Students do not have to limit themselves to alternative heating systems—if they wish, they can also look at ways the homeowner could reduce the amount of heat needed (for example, by insulating their walls). Once the groups are done researching, they can report their results to the class so that everyone can use the same information to do their analysis, or new groups can form so that each member of the group researches a different idea, and the group can work together to use the information each member found.

    Advancing Student Thinking: Students will need to be very careful not to confuse the amount of energy a system uses with the amount it produces, and in their research they will see many ways of describing the energy that is used and produced, such as British thermal units, therms, and Joules. If students struggle with the calculations, it may help if they make a table with column headings like "energy used" and "energy produced" so they can fill in the values they know for each system and calculate the unknown ones.

  - **Modeling Prompt 3B**

    Advancing Student Thinking: Students will need to be very careful not to confuse the amount of energy a system uses with the amount it produces. If students struggle with the calculations, it may help if they make a table with column headings like "energy used" and "energy produced" so they can fill in the values they know for each system and calculate the unknown ones.

- Students can be arranged in groups in advance, they can choose groups, or groups can be determined by using visibly random grouping.

Modeling Prompt 3A

A homeowner wants to replace their old heating system. Energy is measured in kilowatt-hours (kWh). It takes about 11,700 kWh of energy to heat the house for the winter. The current heating system uses natural gas and is 60% efficient, which means that for every 100 kWh of natural gas it uses, it produces 60 kWh of heat. With the homeowner's current system, it costs $975 to heat the house.

Research at least two other options available in your area that this homeowner could replace their heating system with. Assume that natural gas costs $0.05/kWh and electricity costs $0.21/kWh. The house is 2,500 square feet.

The homeowner also has an air conditioner that uses 2,500 kWh of electricity per year and produces 290 kWh of cooling for every 100 kWh it uses. They also have a water heater that uses 4,300 kWh of electricity per year and produces 90 kWh of heat for every 100 kWh it uses. These systems could also be replaced if there is a cheaper option, but it isn't necessary.

1. Which system would you recommend? Make a graph to convince the homeowner to switch to this system.
2. If the homeowner switches to the system you recommend, how long will it take them to save as much money as the new system cost?

Modeling Prompt 3B

A homeowner wants to replace their old heating system. Energy is measured in kilowatt-hours (kWh). It takes about 11,700 kWh of energy to heat the house for the winter. The current heating system uses natural gas and is 60% efficient, which means that for every 100 kWh of natural gas it uses, it produces 60 kWh of heat. With the homeowner's current system, it costs $975 to heat the house. Assume that natural gas costs $0.05/kWh and electricity costs $0.21/kWh.

The homeowner also has an air conditioner that uses 2,500 kWh of electricity per year and produces 290 kWh of cooling for every 100 kWh it uses. They also have a water heater that uses 4,300 kWh of electricity per year and produces 90 kWh of heat for every 100 kWh it uses. These systems could also be replaced if there is a cheaper option, but it isn't necessary.

Here are three other types of heating systems the homeowner could replace their current system with:

- **A new furnace which also runs on natural gas and is more efficient.** For every 100 kWh of natural gas it uses, it produces 95 kWh of heat. This system costs $5,000 to install.
- **A geothermal heat pump.** This system uses electricity instead of natural gas, but it produces 4 kWh of heat for every 1 kWh of electricity it uses. It costs $14,000, but it's very low-maintenance and it also replaces the air conditioner and water heater.
- **A grid-tied solar array.** This system is connected to the electrical grid so that when it generates more energy than the house needs, the extra electricity can be sold back to the grid. Over the whole year, the house will use as much electricity as is sold back, which means the homeowner would basically be heating and cooling the house for free. It costs $16,000 to install.

1. Which system would you recommend? Make a graph to convince the homeowner to switch to this system.
2. How long will it take to save as much money as the new system cost?

Step 4

- Remind students that modeling is a cycle, and that they should evaluate their own models and then refine them, as necessary.
- After sufficient work time, each group or pair should share their solutions with the class. Students could share by presenting to the class, doing a gallery walk, creating a slidedeck, uploading a scan or photo of their work to a shared online space, or by any method that works best for the class.

Step 5

- Provide students time to reflect on their experience with this modeling prompt in their Student Workbook.

RESPONSIVE STRATEGY

As students work in groups, engage them in the Round Robin routine. In this routine, students will each have a turn sharing their thinking, followed by opportunities for group members to ask each other questions and discuss where there are areas of agreement and disagreement.

Round Robin

Modeling Prompt #4: College Characteristics

In this modeling prompt, students will investigate the relationship between two characteristics of colleges. There are two versions of this prompt: 4A and 4B. In 4A, students are given a data sheet with 200 schools and 10 characteristics to choose between. In 4B, students are given a data sheet with 100 schools and 6 characteristics to choose between. Determine in advance which Modeling Prompt (4A or 4B) students will receive, based on the lift-analysis and timing.

Students will need access to technology such as Desmos that will allow them to analyze and display the relationships that interest them.

If desired, more data can be downloaded from the Department of Education website, and data can be curated in a different way. Different categories can be chosen based on the interests of the class, or the range of colleges can be narrowed down to those in the state. Historical data is also available. A few local colleges could be compared across time. If a customized data set is created, modify the introduction to the task accordingly.

Student Task Statement 4A Lift Analysis

Attribute	DQ	QI	SD	AD	M	Avg
Lift	2	1	0	2	2	1.4

Student Task Statement 4B Lift Analysis

Attribute	DQ	QI	SD	AD	M	Avg
Lift	2	1	0	1	2	1.2

Step 1 (*Optional*; review materials as necessary)

- Display and have students reference the Advice on Modeling handout in their Student Workbook. Pass out the Modeling Rubric.
- Facilitate a discussion around modeling. Share some of the following ideas:
  - Modeling prompts are often expressed in words, but unlike word problems, modeling prompts challenge the modeler to make reasonable assumptions, decide what information is important, ask or research for more information if needed, think creatively within constraints, and consider the implications of the model.
  - The process of modeling is cyclical, and it does not end by producing a “correct answer.”
  - A mathematical model expresses a simplified relationship in the real world; models can be rough but still useful.
  - Models can often be refined to represent the real-world relationship more accurately.
  - Responses to modeling prompts can vary widely, but they often contain certain pieces: assumptions, calculations, a mathematical model (stated with an equation or equations, with a graph, with a geometric diagram, or in words), conclusions, and generalizations.
- Provide time for students to ask clarifying questions. Students should have a good-enough understanding of the rubric; deep understanding of the rubric is not needed at this time.

Step 2

- Share with students that the U.S. Department of Education does a survey of colleges every year. Students will use this data to test a conjecture they have about colleges. Tell students that these are some of the characteristics about which there is information:
  - college ownership: public, private non-profit, or private for-profit
  - average SAT scores of admitted students
  - acceptance rate
  - number of enrolled students
  - average cost of attendance
  - median earnings 10 years after graduation

Task Statement 4A has four more categories (highest degree awarded, completion rate, percent of white students, and median ACT score), which can also be included in this discussion if students will be using that task statement.

- Ask students to identify what type of data each of these is (for example, categorical or numerical).
  - "Which of these characteristics might be related to each other?
  - “How are they related?"
- Write down each idea for all to see. Ask students why they think a pair of categories is related in the way they propose, to increase investment in finding the answer. However, students do not need to have well-supported reasons at this time.

If students are having trouble generating ideas, or they do not bring up a wide enough variety of ideas, introduce a few more. Here are some examples:

- The more selective the college is, the higher the average SAT scores of the people who are accepted.
- Private for-profit colleges cost more on average but also have a wider range of costs because they need to make money but also attract lots of students.
- Public schools are larger than private schools.

Once enough ideas have been suggested, ask: “If these relationships exist, how could we see them in the data? How could we present the data to show that they’re true?” Give students quiet think time and then invite them to share one of their ideas with a partner. Then ask students to share their ideas with the class. Ensure that a variety of ways of presenting data are discussed. Here are some examples:

- *The more selective the college is, the higher the average SAT scores of the people who are accepted*: Make a scatter plot of admission rate vs. average SAT score.
- *Private for-profit colleges cost more on average but also have a wider range of costs*: Make a box-and-whisker plot of costs for each type of college.
- *Public schools are larger than private schools*: Make a double bar chart or two way frequency table of school ownership (public or private) vs school size (small, medium, and large, defined by ranges of numbers of enrolled students).

Step 3

- Pass out the pre-determined appropriate blackline master Modeling Prompt #4 (4A or 4B) and Modeling Rubric if it has not already been distributed.
- Have students access the corresponding spreadsheet of college data (4A: https://bit.ly/CollegeData4A and 4B: https://bit.ly/CollegeData4B).
- Students can be arranged in groups in advance, they can choose groups, or groups can be determined by using visibly random grouping.

Modeling Prompt 4A

1. Below are the characteristics of colleges about which there are data. Choose two that might be related. What do you predict the relationship between them is?
   - college ownership: public, private non-profit, or private for-profit
   - average SAT score of admitted students
   - acceptance rate
   - number of enrolled students
   - average cost of attendance per year
   - median earnings 10 years after graduation
   - highest degree awarded

- completion rate
- percent of students who self-identify as white
- median ACT score of admitted students

2. Take a look at the data in the spreadsheet. Before calculating, do some estimation based on the data. Can you tell if your hypothesis seems reasonable?

3. Use an appropriate display to summarize the data.

4. Analyze your graph or table and do any other calculations needed. Do your results confirm your prediction? If not, what do you think explains the results?

5. Here are some headlines. What analysis do you think each article describes? Give some examples of evidence that would support the headline or contradict it.

   - "Students from Selective Colleges Tend to Succeed"
   - "Isolation at Larger Colleges Leads to Higher Dropout Rates"
   - "High Cost of Private Colleges Not Outweighed by Lifetime Earnings"

Modeling Prompt 4B

1. Below are the characteristics of colleges about which there are data. Choose two that might be related. What do you predict the relationship between them is?

   - college ownership: public, private non-profit, or private for-profit
   - average SAT score of admitted students
   - acceptance rate
   - number of enrolled students
   - average cost of attendance per year
   - median earnings 10 years after graduation

2. Take a look at the data in the spreadsheet. Before calculating, do some estimation based on the data. Can you tell if your hypothesis seems reasonable?

3. Use an appropriate display to summarize the data.

4. Analyze your graph or table and do any other calculations needed. Do your results confirm your prediction? If not, what do you think explains the results?

5. Here are some headlines. What analysis do you think each article describes? Give some examples of evidence that would support the headline or contradict it.

   - "Students from Selective Colleges Tend to Succeed"
   - "Isolation at Larger Colleges Leads to Higher Dropout Rates"
   - "High Cost of Private Colleges Not Outweighed by Lifetime Earnings"

Step 4

- Remind students that modeling is a cycle, and that they should evaluate their own models and then refine them, as necessary.

- After sufficient work time, each group or pair should share their solutions with the class. Students could share by presenting to the class, doing a gallery walk, creating a slidedeck, uploading a scan or photo of their work to a shared online space, or by any method that works best for the class.

Step 5

- Provide students time to reflect on their experience with this modeling prompt in their Student Workbook.

TEACHER REFLECTION

Which modeling prompt did you choose for this lesson? What version of the prompt did students receive? How did you make these decisions? Would you make a different choice next time you facilitate this lesson?

Lesson 11: Post-Test Activities

PREPARATION

Lesson Goals	Learning Targets
• Communicate high expectations for all students. • Add and subtract linear expressions.	• I understand the reasoning for and will strive to meet the expectations communicated by my teacher. • I can add and subtract linear expressions.

Lesson Narrative

This lesson, which should occur after the Unit 4 End-of-Unit Assessment, allows for students to reflect on the unit, share feedback, conference with the teacher, and engage in activities that support the work of the upcoming unit.

Gathering student feedback is a powerful and strategic way to learn about students and improve instructional practices. It also creates student and family buy-in and centers students as decision makers and problem solvers in their own learning.

What do you hope to learn about your students during this lesson?

Agenda, Materials, and Preparation

- **Activity 1** (*20 minutes*)
  - End-of-Unit 4 Student Survey (print 1 copy per student)
- **Activity 2** (*25 minutes*)

LESSON

Activity 1: End-of-Unit 4 Student Survey (*20 minutes*)

The End-of-Unit 4 Student Survey is a critical opportunity for teachers to gather low-stakes, non-evaluative feedback to support equity and instructional pedagogy. The survey is also highly beneficial for students as it is designed to encourage self-awareness, self-management, social awareness, relationship skills, and responsible decision making. Provide students a chance to quietly and independently complete this survey after they complete their testing.

One-on-One Conferences

Conducting one-on-one conferences with students, using the surveys as a data point, is encouraged. These conferences can be done as students complete their surveys and are engaging in Activity 2. Potential conference topics include:

- student responses to the daily student reflections
- student response to the end-of-unit student survey (as students finish them)
- executive functioning skills
- student learning contracts
- goal setting and self-evaluation

Activity 2: Area Model (*25 minutes*)

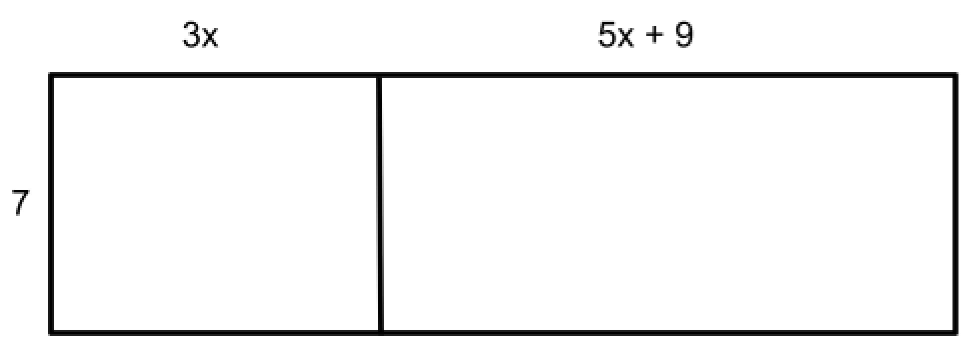
Instructional Routines: Notice and Wonder	
Building On: NC.7.EE.1	**Addressing:** NC.M1.A-APR.1

In this activity, students get a chance to practice adding and subtracting linear expressions from grade 7 in preparation for the work of building functions that occur in upcoming units.

Students can work independently or in pairs to complete this activity. Throughout the activity, students will leverage visual patterns and engage in the *Notice and Wonder* routine as they identify patterns.

Student Task Statement

Rectangle A



1. Write an expression for the area of the entire rectangle A.

   Describe the steps you took and show your work.

2. Double the vertical length of rectangle A to form a new rectangle and call it rectangle B.

   Write an expression that represents the area of rectangle B. Describe the steps you took and show your work.

3. Triple the vertical length of rectangle A to form a new rectangle and call it rectangle C. Write an expression that represents the area of rectangle C.

   Describe the steps you took and show your work.

4. Multiply the vertical length of rectangle A times a factor x to form a new rectangle and call it rectangle D. Write an expression that represents the area of rectangle D.

   Describe the steps you took and show your work.

5. What patterns do you notice in the areas of rectangles A-D? What do you wonder?

6. Add 15 units to each section of rectangle A to form a new rectangle and call it rectangle L. Write an expression that represents the area of rectangle L.

   Describe the steps you took and show your work.

7. Add 50 units to each section of rectangle A to form a new rectangle and call it rectangle M. Write an expression that represents the area of rectangle M.

   Describe the steps you took and show your work.

8. Add x units to each section of rectangle A to form a new rectangle and call it rectangle N. Write an expression that represents the area of rectangle N.

   Describe the steps you took and show your work.

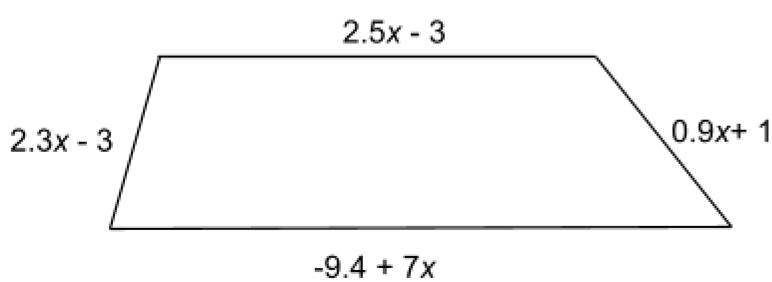
9. What patterns do you notice in the areas of rectangles A, L, M, and N? Describe what you know now.

10. If time is available, choose a few tasks from below to solve:

    a. Write at least two equivalent expressions for the circumference of the circle with a radius of $6x - 3$ meters.

    b. Simplify: $(5x - 8.3) + (2.1x + 3.9) + (4.7 - 3.3x) - (3.2x + 2)$

    c. Find the perimeter of this trapezoid:



TEACHER REFLECTION

As you finish up this unit, reflect on the norms and activities that have supported each student in learning math. List ways you have seen each student grow as a young mathematician throughout this work.

List ways you have seen yourself grow as a teacher.

What will you continue to do and what will you improve upon in Unit 5?