

## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---



INSTRUCTOR SPECIFIC MATERIAL IS INDENTED AND APPEARS IN GREY

**ESTIMATED TIME**

1 hour 20 minutes to 1 hour 40 minutes

**MATERIALS REQUIRED**

Overhead of or electronic display of scatterplots

**BRIEF OVERVIEW**

The series of tasks in this introductory lesson are designed to motivate an initial and informal understanding of concepts related to interpreting scatterplots. This informal understanding is the basis for later formal work with correlation and regression, though those words are not used in this lesson.

**LEARNING GOALS****Students will understand that:**

- Each point on the scatterplot represents a single observation consisting of measurements on two variables.
- An overall downward trend in the data indicates that small values for  $x$  tend to correspond to large values for  $y$ . Larger values for  $x$  tend to correspond to smaller values for  $y$ . Also with a similar understanding of overall upward trends.
- The accuracy in a prediction is related to the variability (scatter) in the data. Variability can be explained by increases in  $x$  or by additional factors having influence on  $y$ .

## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---

**Students will be able to:**

- Interpret the meaning of particular points on the scatterplot.
- Recognize directional trends in the distribution of bivariate data and use these trends to make predictions, without the benefit of formal development of the statistical concepts of association, correlation, and regression.
- Assess the strength of the relationship informally by looking at the degree of scatter.
- Develop plausible explanations for the variability seen in the data.

**INTRODUCTION**

In this series of tasks, you are not formally developing the statistical concepts of correlation or regression. Rather, you are working to build students' ability to see associative trends through the noise of real data and to make judgments in the face of variability. **After** this lesson, students are introduced more formally to the following concepts:

- Form of a relationship as linear or nonlinear,
- Correlation coefficient as a measure of the strength of a linear association, and
- Least squares regression line as a way of describing central tendency of a bivariate distribution (much like the mean describes the central tendency of a univariate distribution).

These concepts are built in later lessons of this module. Within this lesson, students will see opportunities to apply the concepts of 'measures of center' and 'measures of variability'.

**Developmental Math Connections**

Since this is an introduction to working with bivariate data, the learning goals are focused on interpreting scatterplots and motivating an initial and informal understanding of concepts related to distributional thinking with bivariate data that underlies modeling data with a linear function. In future lessons, these ideas become more explicit.

### Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

---

You might wonder why the authors did not begin with a discussion of the Cartesian plane or an exercise in which students construct a scatterplot, as you might in an elementary algebra course. The rationale is multifold. First, in Statway is designed for a more heterogeneous group of students. Therefore, you need to find ways to level the playing field so that some students are not bored and others are not lost. Of course, this level playing field means that all students must be able to bring something to the task that puts them into play, so to speak. Second, the authors build some of these skills directly into the lesson. For example, students plot points in Task 2, but the coordinates of points is connected to using a scatterplot. Point plotting and coordinates is in the service of the learning goal of interpreting trends in bivariate data. Third, research suggests that students need to struggle with meaningful tasks before a careful explication of concepts helps them construct deeply held understandings. The lesson is constructed to provide an opportunity for productive struggle, followed by carefully constructed tasks that the authors think are sufficient for all students to begin to make sense of the concepts at hand.

### Lesson Structure

This lesson has the following components:

- Introduction to the context of the lesson (10 minutes).
- Part 1: Students work on a rich task, wrap-up, and transition to Part 2 (20–25 minutes).
- Part 2: Scaffolded conceptual tasks (via discussion/group work/lecture) and wrap-up (40–50 minutes).
- Homework (done outside of class) (5 minutes in class for directions).

### Introduction to the Context of the Lesson

To orient students to the data set, distribute the following handout and go over the introductory material before giving them a few minutes to choose the most nutritious cereal.

## Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships



STUDENT MATERIAL IS NOT INDENTED AND APPEARS IN BLACK

## INTRODUCTION

In the United States, the Food and Drug Administration requires that most prepared food be labeled with nutrition information. Today part of the data set for the lesson was gathered from the list of required nutrition facts that appear on boxes of breakfast cereals.

Below are the nutrition facts for three cereals. There is a lot of information given. An entire math lesson could probably be designed around reading and understanding the information given on cereal labels. For now though, you will look over the labels fairly quickly, like you might do in the grocery store. Use your initial impressions to pick the cereal that you would rate as the most nutritious (the best for your health). You have a few minutes to do this.

<b>Nutrition Facts</b> Serving Size 1.0 cup (1 NLEA serving) <hr/> <b>Amount Per Serving</b> <b>Calories</b> 117 <b>Calories from Fat</b> 9 <hr/> <b>% Daily Value *</b> <b>Total Fat</b> 1g                      1% Saturated Fat 0g                      1% Polyunsaturated Fat 0g Monounsaturated Fat 1g <b>Cholesterol</b> 0mg                      0% <b>Sodium</b> 171mg                      7% <b>Total Carbohydrate</b> 26g                      9% Dietary Fiber 1g                      3% Sugars 14g <b>Protein</b> 1g <hr/> Vitamin A 0%    •    Vitamin C 10% Calcium 10%    •    Iron 30% <hr/> * Percent Daily Values are based on a 2,000 calorie diet. Your daily value may be higher or lower depending on your calorie needs: <table> <tr> <td></td> <td>Calories: 2,000</td> <td>2,500</td> </tr> <tr> <td>Total Fat</td> <td>Less than 65g</td> <td>80g</td> </tr> <tr> <td>Sat Fat</td> <td>Less than 20g</td> <td>25g</td> </tr> <tr> <td>Cholesterol</td> <td>Less than 300mg</td> <td>300mg</td> </tr> <tr> <td>Sodium</td> <td>Less than 2,400mg</td> <td>2,400mg</td> </tr> <tr> <td>Total Carbohydrate</td> <td>300g</td> <td>375g</td> </tr> <tr> <td>Dietary Fiber</td> <td>25g</td> <td>30g</td> </tr> </table> <hr/> Calories per gram: Fat 9 • Carbohydrate 4 • Protein 4		Calories: 2,000	2,500	Total Fat	Less than 65g	80g	Sat Fat	Less than 20g	25g	Cholesterol	Less than 300mg	300mg	Sodium	Less than 2,400mg	2,400mg	Total Carbohydrate	300g	375g	Dietary Fiber	25g	30g	<b>Nutrition Facts</b> Serving Size 1.0 cup (1 NLEA serving) <hr/> <b>Amount Per Serving</b> <b>Calories</b> 101 <b>Calories from Fat</b> 1 <hr/> <b>% Daily Value *</b> <b>Total Fat</b> 0g                      0% Saturated Fat 0g                      0% Polyunsaturated Fat 0g Monounsaturated Fat 0g <b>Cholesterol</b> 0mg                      0% <b>Sodium</b> 202mg                      8% <b>Total Carbohydrate</b> 24g                      8% Dietary Fiber 1g                      3% Sugars 3g <b>Protein</b> 2g <hr/> Vitamin A 0%    •    Vitamin C 10% Calcium 0%    •    Iron 54% <hr/> * Percent Daily Values are based on a 2,000 calorie diet. Your daily value may be higher or lower depending on your calorie needs: <table> <tr> <td></td> <td>Calories: 2,000</td> <td>2,500</td> </tr> <tr> <td>Total Fat</td> <td>Less than 65g</td> <td>80g</td> </tr> <tr> <td>Sat Fat</td> <td>Less than 20g</td> <td>25g</td> </tr> <tr> <td>Cholesterol</td> <td>Less than 300mg</td> <td>300mg</td> </tr> <tr> <td>Sodium</td> <td>Less than 2,400mg</td> <td>2,400mg</td> </tr> <tr> <td>Total Carbohydrate</td> <td>300g</td> <td>375g</td> </tr> <tr> <td>Dietary Fiber</td> <td>25g</td> <td>30g</td> </tr> </table> <hr/> Calories per gram: Fat 9 • Carbohydrate 4 • Protein 4		Calories: 2,000	2,500	Total Fat	Less than 65g	80g	Sat Fat	Less than 20g	25g	Cholesterol	Less than 300mg	300mg	Sodium	Less than 2,400mg	2,400mg	Total Carbohydrate	300g	375g	Dietary Fiber	25g	30g	<b>Nutrition Facts</b> Serving Size 1.0 cup (1 NLEA serving) <hr/> <b>Amount Per Serving</b> <b>Calories</b> 111 <b>Calories from Fat</b> 16 <hr/> <b>% Daily Value *</b> <b>Total Fat</b> 2g                      3% Saturated Fat 0g                      2% Polyunsaturated Fat 1g Monounsaturated Fat 1g <b>Cholesterol</b> 0mg                      0% <b>Sodium</b> 213mg                      9% <b>Total Carbohydrate</b> 22g                      7% Dietary Fiber 4g                      14% Sugars 1g <b>Protein</b> 4g <hr/> Vitamin A 0%    •    Vitamin C 10% Calcium 12%    •    Iron 69% <hr/> * Percent Daily Values are based on a 2,000 calorie diet. Your daily value may be higher or lower depending on your calorie needs: <table> <tr> <td></td> <td>Calories: 2,000</td> <td>2,500</td> </tr> <tr> <td>Total Fat</td> <td>Less than 65g</td> <td>80g</td> </tr> <tr> <td>Sat Fat</td> <td>Less than 20g</td> <td>25g</td> </tr> <tr> <td>Cholesterol</td> <td>Less than 300mg</td> <td>300mg</td> </tr> <tr> <td>Sodium</td> <td>Less than 2,400mg</td> <td>2,400mg</td> </tr> <tr> <td>Total Carbohydrate</td> <td>300g</td> <td>375g</td> </tr> <tr> <td>Dietary Fiber</td> <td>25g</td> <td>30g</td> </tr> </table> <hr/> Calories per gram: Fat 9 • Carbohydrate 4 • Protein 4		Calories: 2,000	2,500	Total Fat	Less than 65g	80g	Sat Fat	Less than 20g	25g	Cholesterol	Less than 300mg	300mg	Sodium	Less than 2,400mg	2,400mg	Total Carbohydrate	300g	375g	Dietary Fiber	25g	30g
	Calories: 2,000	2,500																																																															
Total Fat	Less than 65g	80g																																																															
Sat Fat	Less than 20g	25g																																																															
Cholesterol	Less than 300mg	300mg																																																															
Sodium	Less than 2,400mg	2,400mg																																																															
Total Carbohydrate	300g	375g																																																															
Dietary Fiber	25g	30g																																																															
	Calories: 2,000	2,500																																																															
Total Fat	Less than 65g	80g																																																															
Sat Fat	Less than 20g	25g																																																															
Cholesterol	Less than 300mg	300mg																																																															
Sodium	Less than 2,400mg	2,400mg																																																															
Total Carbohydrate	300g	375g																																																															
Dietary Fiber	25g	30g																																																															
	Calories: 2,000	2,500																																																															
Total Fat	Less than 65g	80g																																																															
Sat Fat	Less than 20g	25g																																																															
Cholesterol	Less than 300mg	300mg																																																															
Sodium	Less than 2,400mg	2,400mg																																																															
Total Carbohydrate	300g	375g																																																															
Dietary Fiber	25g	30g																																																															

**Note:** These cereals are [from left to right] General Mills Cocoa Puffs, Kellogg's Corn Flakes, and General Mills Cheerios. Do not tell students the name of the cereals until

### Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

---

after they have made their choices because the names might divert their attention from using the data to make their decision.

### WRAP-UP

After giving students a few minutes to make a decision, take a quick tally to see which cereal students chose as most nutritious. Then call on a few students to describe which ingredients they focused on. The goal in this short interaction with the class is to highlight the following three points to get students thinking about how ingredient amounts might be associated with nutritional ratings:

- Different people might devise different ways to rate the nutrition of cereals by focusing on different ingredients.
- You could use one ingredient or more than one ingredient to determine a nutrition rating.
- In a rating system, there is an association between the ingredient and the nutrition rating. For example, large amounts of sugar might give lower ratings. Large amounts of fiber might give higher ratings.

### INTRODUCTION TO PART 1

Give students about 3 minutes to work on these questions alone and then in small groups for 5–10 minutes (depending on your sense of whether productive conversations are occurring).

Next Steps provides students the opportunity to struggle with important ideas (like interpreting scatterplots and seeing patterns that relate to a question at hand). So, at this point in the lesson, you do not need to guide students to discover canonical ideas, such as correlation, or even correct their misconceptions or fix their errors. This is an informal introduction to distributional thinking with bivariate data. While students work, listen to how students are reasoning as they discuss the task. In the wrap-up, you have the opportunity to talk with students in general terms about making predictions and using a visual sense of the variability in the data to determine which ingredient is a more accurate predictor of ratings. In the wrap-up, you can refer to what you observed as students worked, giving praise and noting interesting aspects of their conversations that are relevant to learning goals for the lesson.

## Lesson 3.1.1

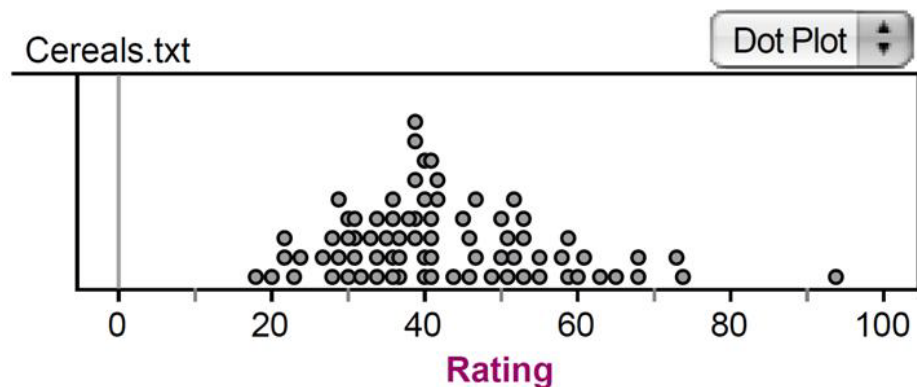
## Introduction to Scatterplots and Bivariate Relationships

Now transition to the *Consumer Reports* ratings by using the following information on the student handout to discuss *Consumer Reports*.

## NEXT STEPS – PART 1

- 1 Instead of making up your own rating system, you are going to investigate the *Consumer Reports* nutritional ratings for 77 breakfast cereals. *Consumer Reports* is published by a nonprofit organization called the Consumers Union, whose mission it is to work for a fair, just, and safe marketplace for all consumers and to empower consumers to protect themselves. *Consumer Reports* rates products based on its own criteria and testing. It prides itself on producing objective results. *Consumer Reports* maintain its objectivity by not allowing advertising within their publications and not allowing use of their results for commercial gain. (Retrieved from [www.consumerreports.org/cro/aboutus/mission/overview/index.htm](http://www.consumerreports.org/cro/aboutus/mission/overview/index.htm))

*Consumer Reports* uses a rating system with a scale of 0 to 100. Here is the distribution of *Consumer Reports* ratings for 77 cereals:



- A What does each dot represent in this distribution?

**Answer:** 1 cereal.

- B For this distribution, what seems to be an average rating?

**Answer:** Approximately 40 to 45.

- C How would you describe the variability in ratings? Find the range of the ratings, and guess the standard deviation.

## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---

**Answer:** Most ratings are between 20 and 60. The range is between 18 and 94, the standard deviation is approximately 10 to 15.

- D How would you describe the shape of this distribution? What does the shape suggest about the rating system?

**Answer:** Skewed right. The shape of the dotplot tells us nothing about the rating system.

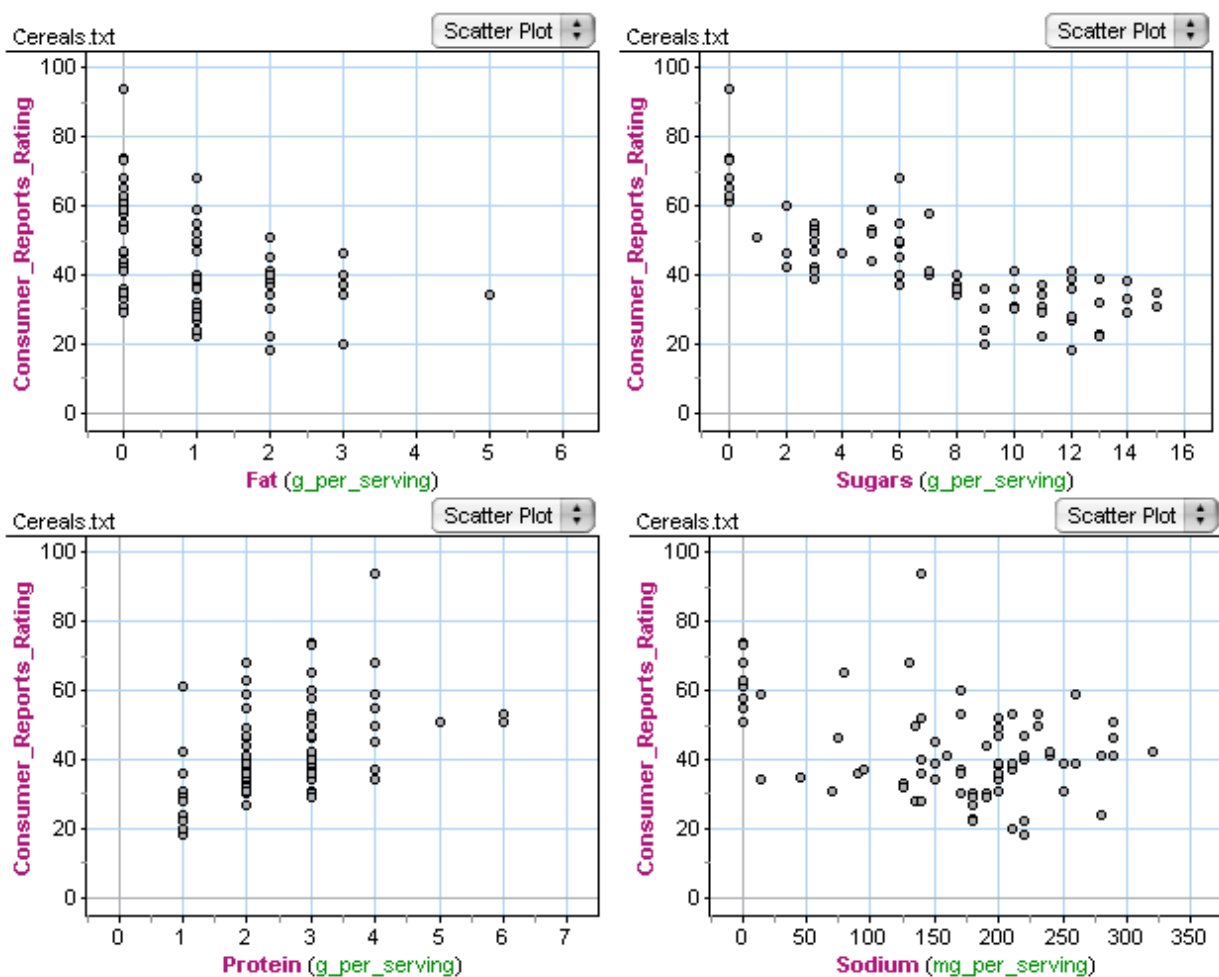
**Note:** Students should be able to describe shape, center, and spread of this distribution based on what they have learned in Module 2. Answers can vary. Reasonable responses include the following: The dots in the distribution are cereals; The average rating is about 40 [actual mean is 44.0, median 41.0]; Ratings fall between about 18 and 94 on a scale of 0 to 100, with a reasonable estimate for the standard deviation being 10 rating points [actual standard deviation is 13.8, Q1 34, Q3 52.9].

## Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

What you cannot tell from the dotplot is how the cereal ingredients (such as sugar or fat) are related to the ratings. You need a new type of graph, called a scatterplot, to investigate how two variables relate to each other. The scatterplots below show the amount of an ingredient in a serving of cereal and the *Consumer Reports* rating for 77 breakfast cereals.

In these graphs, the rating is now shown in the vertical direction



The *Consumer Reports* rating formula is not made public. You do not know which ingredients are used in its rating formula. In this lesson, you will try to “crack their code” to identify the more important ingredients for their rating. We will use the data to figure out which ingredients *Consumer Reports* may, or may not, use in their rating formula. The only clues you have are these scatterplots.



## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---

The first step in this investigation is to answer the following two questions AND write down enough of your reasoning that someone can follow your thinking.

*Two new cereals are being rated by Consumer Reports. Cereal A has 10.5 grams of sugar in a serving and Cereal B has 2.5 grams of protein in a serving.*

- 2 Based on the data shown, predict the *Consumer Reports* rating for the two cereals.

**Answer:** Cereal A: Between 20 and 40 would be reasonable. Draw a vertical line segment at  $x = 10.5$ ,  $20 < y < 40$  to visually represent this reasonable range of predictions. Cereal B: between 25 and 70 would be reasonable. Represent this range with a vertical line segment.

- 3 For which cereal do you think your prediction is probably more accurate (more likely to be closer to the actual Consumer Report rating)? Why?

**Answer:** For Cereal A, there is a narrower range of reasonable predictions for the *Consumer Reports* rating (approximately 20 to 40), so the amount of sugar is a fairly accurate predictor of the rating. For Cereal B, there is a wider range of reasonable predictions for the *Consumer Reports* rating (approximately 25 to 70), so the amount of protein is not as accurate a predictor of the rating. There are other factors influencing the rating besides the amount of protein. So, it is more likely that your prediction of the rating for Cereal A is closer to the actual rating than your prediction for Cereal B.

**Direct Instruction About Statistical Concepts**

You will need to project the scatterplots for this discussion. To see if students are using the patterns in the data to make predictions, ask them to determine if the following predictions for ratings are reasonable or unreasonable (perhaps ask students to show a thumbs up for *reasonable prediction* and thumbs down for *unreasonable prediction*):  
Cereal A: 10, 30, 60; Cereal B: 10, 30, 60.

Plot each of these predictions on the scatterplot and highlight how the prediction fits the pattern in the data or deviates from the pattern.

Cereal A: 10 (not reasonable), 30 (reasonable), 60 (not reasonable); Cereal B: 10 (not reasonable), 30 (reasonable), 60 (reasonable).

Discuss the following questions through a brief mini-lecture or class discussion:

## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---

- What is a range of reasonable predictions for ratings of Cereal A? of Cereal B?
- Which ingredient, sugar or protein, is a more accurate predictor of *Consumer Reports* ratings?

The next segment of the lesson, Try These, is designed with more explicit attention to developing the skills and understandings described in the learning outcomes. It is in Part II that you make connections as well correct errors and address misconceptions as appropriate.

**TRY THESE****Reading and Interpreting Scatterplots**

You are going to take a short detour from your investigation into which ingredients are the best predictors of *Consumer Reports* ratings. Here, you will work on interpreting scatterplots just to make sure everyone is comfortable with reading this type of graph.

- 4 Captain Crunch has the lowest *Consumer Reports* rating of the 77 cereals in the data set. How much fat is in a serving of Captain Crunch?

**Answer:** 2 grams.

- 5 In this set of 77 cereals, Product 19 has the most sodium in a serving. What is the rating for Product 19?

**Answer:** about 42.

- 6 All-Bran Extra Fiber is the cereal with the highest rating. How much sugar, fat, and sodium are in a serving of All-Bran Extra Fiber?

**Answer:** 0 grams sugar, 0 grams fat, about 140grams sodium.

**Direct Instruction About Statistical Concepts**

If students have been working in groups, you will have some sense of where they had difficulty. For the subsequent whole-class discussion, address areas of difficulty and answer questions in the context of the wrap-up questions. You will not have time to go over the answers to the scaffolded questions.

### Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

---

To bring closure, ask students to discuss the following questions or deliver a short mini-lecture that answers these questions:

When a statistician reads a scatterplot, he or she asks herself two questions:

- (1) Who or what is described by the data? (i.e., What does a dot represent?) and
- (2) What measurements were made? (i.e., What are the variables?) Pick a scatterplot, and answer these two questions.

### INTRODUCTION TO PART 2

Students will be exploring patterns and trends in the data.

### NEXT STEPS – PART 2

#### Seeing Patterns and Relationships in Scatterplots

Now you will continue your detective work with *Consumer Reports* ratings. Try to identify ingredients that are good predictors of ratings and ingredients that are not. More importantly, focus on how patterns in the data are related to identifying ingredients that are good predictors.

**Note:** If students are working in groups, for most of these activities, the authors suggest that you intervene if students have answers that are wildly off base otherwise not. The goal is to foster distributional thinking with bivariate data. So, when you intervene, refrain from correcting; instead get students to talk about what they are seeing. Make observations or ask questions to nudge them in the right direction. Remember that in the wrap-up, you will provide direct instruction relative to the learning goals for the lesson, so you do not have to fix everything during group work.

- 7 There are four cereals that have 3 grams of fat in a serving. Estimate the ratings for these four cereals. What might explain the variability in the ratings?

**Answer:** Estimates are about 20, 34, 37, 46. Intervene and correct if students are misreading the scatterplots. Variability is explained by the impact of other ingredients on the ratings.

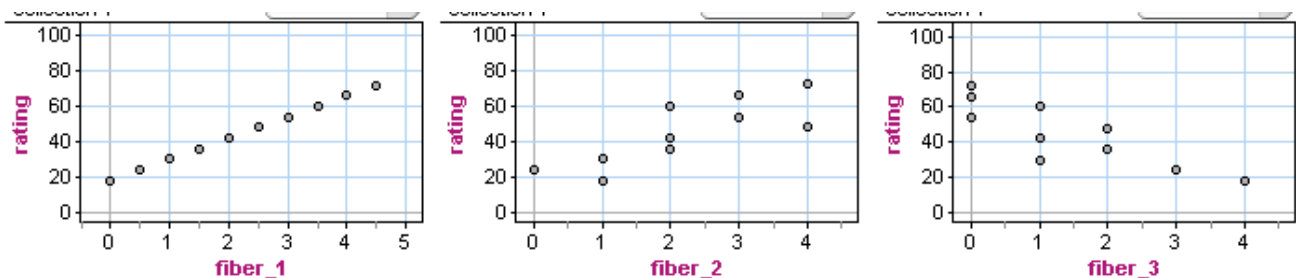
## Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

- 8 Imagine changing the recipe for a cereal that has 0 grams of fat in a serving and a rating of 60. Increase the amount of fat to 3 grams in a serving. Do you think the rating will probably increase or decrease or remain about the same? Or do you think that it is impossible to use the scatterplot to predict the impact of this change on the rating? How does the pattern in the data support your decision?

**Answer:** The rating will probably decrease. You see a downward trend in the data that suggests that larger amounts of fat in a serving tend to be associated with lower ratings. If students think that it is impossible to predict how the rating will change, ask them to say more about why this is so. The goal is to get students to begin to develop distributional thinking for bivariate data, so nudge them to think about trends and patterns. It is reasonable to be unsure about how the rating will change because of the noise in the data. One pattern they could describe to support the “cannot tell” response is the large amount of variability in ratings for cereals with the same amount of fat.

- 9 Think about how the amount of fiber in a cereal might relate to the *Consumer Reports* rating. Here are three scatterplots with make-believe data from 10 imaginary cereals. Which scatterplot do you think displays a pattern similar to what you may see in the actual data? Why?



**Answer:** The middle graph. Intervene, if students pick another graph. Get them to talk about what they are seeing. Do they know that fiber is a good thing? Why would you expect some variability in ratings for cereals with the same amount of fiber?

### Direct Instruction About Statistical Concepts

In this wrap-up, return to the issue of cracking the code on the *Consumer Reports* ratings by determining which ingredients appear to be good predictors of ratings. The real goal is to focus on issues related to building distributional thinking in this new setting of bivariate data. By the end of this wrap-up, the following should be clear to students:

## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---

- Upward and downward trends can help them make predictions even with very noisy data.
- The more scatter (variability) in the data, the less accurate their predictions probably are.
- When there is a lot of variability in the ratings for a fixed amount of ingredient, the ingredient is not a good predictor of the rating. Other factors are affecting the rating.

Compare and contrast fat-ratings and sugar-ratings scatterplots to address issues related to the previous activities:

Both scatterplots show a downward trend that indicates that cereals with smaller amounts of fat (or sugar) tend to have higher ratings and cereals with larger amounts of fat (or sugar) tend to have lower ratings. Draw a summarizing line with a negative slope or a highlight a downward diagonal region to highlight the trend. The downward trends indicate that increasing the amount of fat (or sugar) in a serving will tend to decrease the *Consumer Reports* rating.

The downward trends suggest that both fat and sugar are used in the *Consumer Reports* rating, but sugar is a better predictor of the ratings than fat because there is less variability in the ratings for similar cereals. Here you are reasoning in the same way that you did previously. For example, when you look at cereals with 1 gram of fat in a serving, you see a wide range of ratings falling between 20 and 70. You tend to see a similar wide range of ratings when you compare cereals with the same amount of fat. If you compare cereals with the same amount of sugar, you usually see less variability in their ratings. For example, cereals with 6 grams of sugar in a serving have ratings that differ, but the range of variability is less (about 35 to 65). The variability that we have been discussing is due to the impact of other ingredients on the rating.

### ***Discussion of the Fiber-Ratings Relationship***

Poll student responses to fiber-ratings problem (Question 9). Look at actual fiber data and compare and contrast it with the hypothetical data pattern that the majority chose.

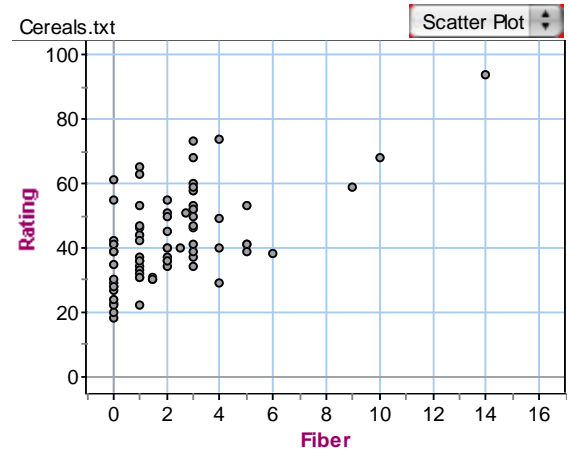
In what ways does the scatterplot of the real data look like what they expected? You might have expected the upward trend in the data since smaller amounts of fiber

## Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

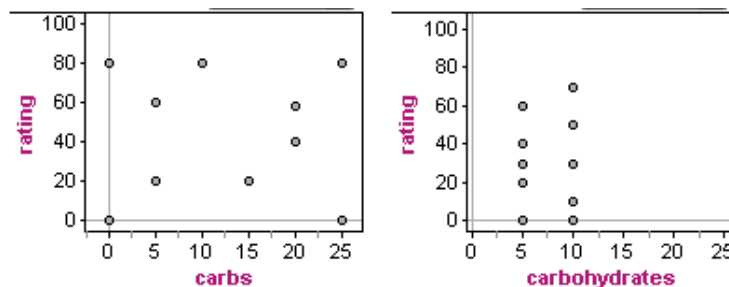
would probably result in lower ratings and larger amounts of fiber in higher ratings. Highlight this upward trend with a line or a diagonal region.

In what ways is this graph surprising? Again, you see a surprising amount of variability in the ratings for cereals with the same amount of fiber. This suggests that fiber is used in the rating formula but does not have as strong an impact as sugar. You might also be surprised that there are only a few cereals with more than 6 grams of fiber in a serving. These cereals have pretty high ratings.



### *Discussion of Graphs Showing No Relationship Between Carbohydrates and Ratings*

Here are two graphs that illustrate no relationship between carbohydrates and ratings. In the graph on the left, there is no discernable pattern that can be used to make predictions. In the graph on the right, there is a vertical pattern, but knowing the amount of carbohydrate does not help predict the rating. The vertical pattern is simply caused by all the dots being at two values of carbohydrates.



## Lesson 3.1.1

**Introduction to Scatterplots and Bivariate Relationships**

---

**TAKE IT HOME**

- 1 Summarize what you feel you learned today.
- 2 The average *Consumer Reports* rating for these 77 cereals is 44. What is the largest amount of sugar per serving in a cereal that has above average ratings?

**Answer:** 7 grams per serving.

- 3 Which is a better predictor of the Consumer Report ratings: sugar or sodium? Explain how the scatterplots support your answer.

**Answer:** Sugar is a better predictor of ratings. There is a clear downward trend that makes it easier to predict the ratings based on a given amount of sugar. The sodium-ratings scatterplot has a lot of variability in ratings for cereals with similar amounts of sodium.

- 4 A friend says that she only pays attention to sugar amounts, even though she is also concerned by fat. Her reasoning is that low levels of sugar signal that the food also has low amounts of fat. Similarly, high levels of sugar signal that the food also has high amounts of fat. The scatterplot shows the sugar and fat content of cereals. Does this appear to be true for breakfast cereals? Explain how the scatterplot supports your answer.

**Answer:** This is not true for the breakfast cereals in this data set. The pattern described is an upward trend, but this graph does not have an upward trend. In this graph the cereals with highest amounts of sugar [14–16 grams in a serving], have low amounts of fat (0–1 grams in a serving.)

### Lesson 3.1.1

## Introduction to Scatterplots and Bivariate Relationships

---

+++++

This lesson is part of STATWAY™, A Pathway Through College Statistics, which is a product of a Carnegie Networked Improvement Community that seeks to advance student success. Version 1.0, A Pathway Through Statistics, Statway™ was created by the Charles A. Dana Center at the University of Texas at Austin under sponsorship of the Carnegie Foundation for the Advancement of Teaching. This version 1.5 and all subsequent versions, result from the continuous improvement efforts of the Carnegie Networked Improvement Community. The network brings together community college faculty and staff, designers, researchers and developers. It is an open-resource research and development community that seeks to harvest the wisdom of its diverse participants in systematic and disciplined inquiries to improve developmental mathematics instruction. For more information on the Statway Networked Improvement Community, please visit [carnegiefoundation.org](http://carnegiefoundation.org). For the most recent version of instructional materials, visit [Statway.org/kernel](http://Statway.org/kernel).

+++++

STATWAY™ and the Carnegie Foundation logo are trademarks of the Carnegie Foundation for the Advancement of Teaching. A Pathway Through College Statistics may be used as provided in the CC BY license, but neither the Statway trademark nor the Carnegie Foundation logo may be used without the prior written consent of the Carnegie Foundation.