OVERVIEW

The main focus of this activity is comparing groups—comparing backpack weights for students in different grades and comparing backpack weights for boys and girls. The level of sophistication that students use for these comparisons will vary by grade level, but all comparisons should include some descriptions and use of centers, such as means or center clumps. Later grades may also focus on differences in spread among the groups.

The activity Is Your Backpack Too Heavy for You? extends the themes of this activity and could be used in conjunction with it.

For grades 4–5, this activity also provides a good opportunity to distinguish between categorical and numerical data. Students should be able to identify the categorical attributes (*Name, Gender, Grade*) from the numerical attributes (*BodyWeight, PackWeight*). If you use this activity to introduce TinkerPlots, you can have students explore how TinkerPlots uses color for different types of attributes and what happens when you separate categorical versus numerical attributes.

For all grades, students should explore different ways of separating, ordering, and stacking the data, creating a variety of graphical representations. They should observe how different plots highlight different aspects of the data and choose a plot that helps them tell the story they want to tell.

Activity Time: One class period

Objectives

- Compare related sets of data.
- Make observations about differences between groups.
- Distinguish between categorical and numerical data.
- Describe the shape and other important aspects of numerical data.
- Justify conclusions based on data.
- Represent data with graphs.

Common Core Standards Addressed

Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.

Grade 6, Statistics and Probability Standard 2

Summarize numerical data sets in relation to their context.

Grade 6, Statistics and Probability Standard 5

Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.

Grade 6, Statistics and Probability Standard 2

Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge variation in estimates or predications.

Grade 7, Statistics and Probability Standard 2

Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations.

Grade 7, Statistics and Probability Standard 4

Prerequisites

- Students need to know that TinkerPlots collects data in *data cards* and that *plots* allow you to display and organize data. These are introduced in the movie "TinkerPlots Basics."

Materials

- Who Has the Heaviest Backpacks? worksheet (one per student)
- **Heaviest Backpacks.tp**

Lesson Plan

Think About It (10 minutes)

Hand out the Who Has the Heaviest Backpacks? worksheet.

Encourage students to work in pairs or small groups to write answers for the Think About It questions, or have students write individual answers and then discuss them in groups. Involving students in group discussions will foster communication, help make apparent common expectations about the data and questions, and illuminate alternative ideas.

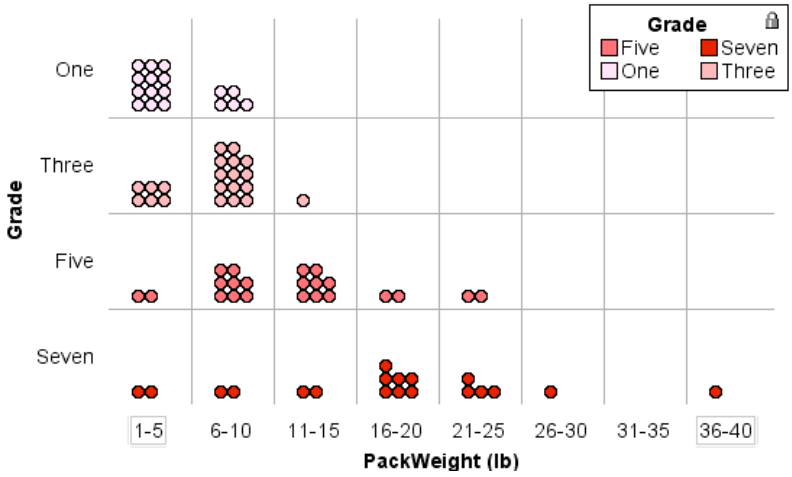
Note: Step 1 uses the figure of speech "on average." Although students in grades 6–8 may immediately think "mathematical average," the question is not necessarily asking for the mean. Encourage students to think about what backpack weight or weights are "typical." Students are likely to use their own backpacks as a way to estimate; if so, be sure to have them adjust their estimates to include the whole group in question, which includes first-, third-, fifth-, and seventh-grade students.

Plot and Investigate (15 minutes)

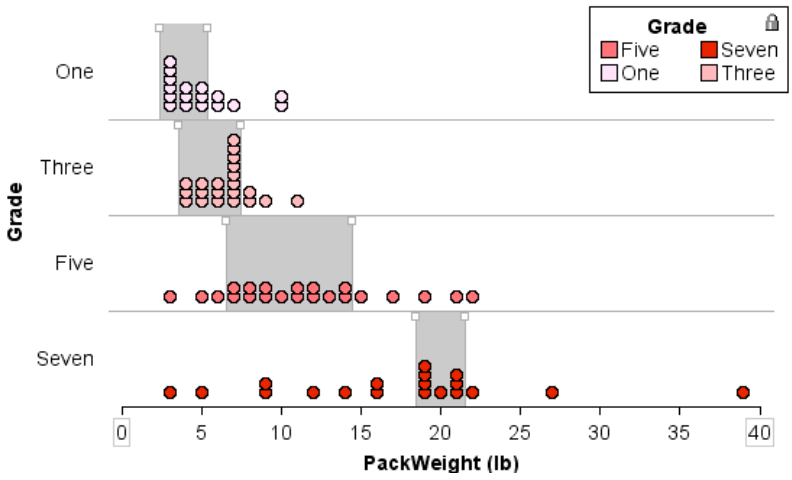
Have students move to computers and open the document **Heaviest Backpacks.tp** to explore backpack weight data.

In exploring Step 6, students in grades 4–5 should, at the least, use a plot that is ordered by *PackWeight* and colored by *Grade*. Students should then note that first-graders (light-pink) tend to be on the far left (lower weights); third-graders (pink) tend to be in the middle left; fifth-graders (red) tend to be in the middle right; and seventh-graders (dark red) tend to be on the far right (higher weights).

Students may separate the case icons into bins, as shown on the next page. Many students will want to compare the counts in common bins: "There are more students that carry backpacks above 16 pounds in grade 7 than in any other grade. So students in grade 7 carry heavier backpacks." However, because the data do not include the same total number of students in each grade, this is not a valid argument. Students should instead think in terms of fractions or percents ("68% of students in grade 7 carry backpacks over 15 pounds, but only 18% of students in grade 5 and 0% in grades 1 and 3 carry backpacks over 15 pounds. So students in grade 7 carry heavier backpacks"). Alternatively, students could focus on center clumps ("Grade 1 cluster around 1–5 pounds. Grade 3 cluster around 6–10 pounds. Grade 5 cluster around 6–15 pounds. And grade 7 cluster around 16–25 pounds.")



Some students, especially those in grades 6–8, will fully separate both *Grade* and *PackWeight*. They might use various ways to highlight and argue why the data show students in higher grades tend to carry heavier backpacks. The plot on the next page uses dividers to highlight center clusters. Here, the clusters in the gray divisions show an increase as the grades increase.



Depending on their proficiency with TinkerPlots, students in grades 4–5 might include range hats, the mode, or the median to further support their answer. Students in grades 6–8 might use box plots or the mean.

Commend students who try other plots that clearly support their conclusions, including students who are able to construct plots that support alternative conclusions. Encourage older students to refine their conclusions by using precise statements, such as "50% of seventh-graders have a backpack between 12 and 21 pounds."

Wrap-Up (10 minutes)

Have a couple of students with different kinds of plots present them to the class. You can do this by having students walk around from computer to computer, or, if you have remote connection, by displaying their computer screens on the overhead display. Discuss how different plots help to answer the questions posed in Steps 7 and 9.

Extensions (optional)

1. Have students conduct their own study, collect data about backpacks at your school, and analyze the results. Compare the results from your school with the results for the data in **Heaviest Backpacks.tp.** You may want to do this along with a talk about *samples* and *populations*, discussing whether students think data from any one school are likely to be representative of schools across the country. Students could even plan ways to collect more representative data (for example, by collecting data at several different schools or by gathering statewide or nationwide data via the Internet).
2. Hold a discussion about ways students could lighten their backpacks or about the proper way to carry a backpack to minimize stress. See the links in the TinkerPlots Online Resource Center for more information. If students conduct their own study, they may want to collect additional data about the way each student carries his or her backpack and whether each student has experienced back pain.
3. Discuss *outliers.* For grades 4–5, ask students what they think an outlier is, and then have them identify data values that might be outliers. (Faith, the seventh-grader with the 39-pound backpack, is an obvious outlier.) Rather than focusing too much on the specific definition of an outlier, encourage students to consider whether the outliers' values are

correct or not. Sometimes values that are extremely far away are errors, while other times they are actual values and thus of particular interest. More typically, as is the case here, there is no way to tell for certain.

Students in grades 6–8 who used the mean or median to summarize the data could delete the outliers from the collection or change their values and see how this affects the average.

4. Ask questions that challenge students to predict cases that were not sampled. For example, "Judy is a seventh-grader. She was absent on the day these data were collected. If you had to guess the weight of her backpack, what would you guess?" You can also have students make predictions for whole groups: "Suppose we measure the backpack weights of second-graders at the same school. What would the average weight be? Make up a reasonable collection of 20 cases for the second-graders." Students can add these cases to the data set.

ANSWERS

1. Answers will vary depending on grade level and estimation skills. Many students will want to give a range for an answer (for example, "5 to 10 pounds"), and this is perfectly acceptable; there is no reason to force students into giving a single number. As guideposts for answers, the data in **Heaviest Backpacks.tp** have a mean of 10.2 pounds and a median of 8 pounds. The middle 50% of the values range from about 5 to 14 pounds.

   Students may have trouble estimating weights in pounds, so you may want to have some reference weights available, such as 3-, 5-, and 10-pound weights from your gymnasium or 5- and 10-pound bags of sugar.

2. Students will likely guess that students in higher grades carry heavier backpacks. Explanations may include heavier books in higher grades, more books or homework in higher grades, or additional items such as sports equipment.

3. Some students might guess that girls carry heavier backpacks because they are (stereotypically) more studious; some might guess that boys carry heavier backpacks because they (stereotypically) want to show how strong they are, or carry extra items such as sports equipment; and some might think the weights are the same because both boys and girls are in the same classes and have the same books.

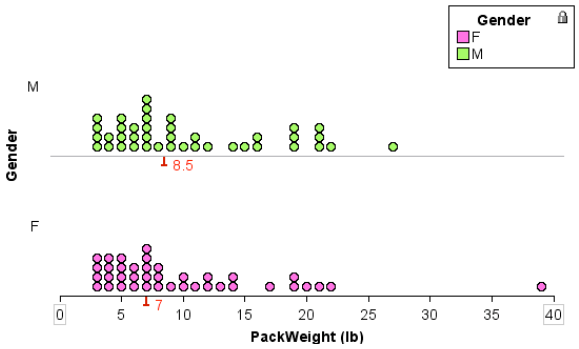
5. 79 students. Students can find this answer from the upper-right corner of the data cards, by counting the circles in the plot, or by adding counts to the plot.

7. Yes, students in higher grades tend to carry heavier backpacks than students in lower grades. Explanations must include how the plot supports their answers.

9. Answers will vary depending on the plots and statistics used. Again, students must explain how the plot supports their answers.

   A simple plot ordered by *PackWeight* and colored by *Gender* may show no noticeable difference between boys and girls. Even a more sophisticated plot, such as a stacked dot plot (see below), may show little or no difference. Students might explain that these plots show about the same number of boys and girls with heavy and light backpacks, or that the clusters are located in about the same places.

If older or advanced students use hat plots, the median, or the mean, they may notice that boys carry slightly heavier backpacks than girls. For the plot below, a student might say, "The median for boys is 8.5 and the median for girls is 7, so boys carry heavier backpacks." Even so, some students may feel that the difference is not significant enough and still say that boys and girls tend to carry about the same amount of weight in their backpacks. (In fact, this is what a statistician would likely say as well.)



Note: So that students are able to create two plots in their document—one for Step 5 and one for Step 6, you may need to show them how to use the lock icon **Key** to lock the plot color (as shown in all of the plots above).