

OVERVIEW

In this activity, students explore adding or subtracting a constant as a way to create a group difference, and explore three contexts that can be modeled using this additive-shift metaphor.

To complete the worksheet, students should know how to create a plot by dragging attributes to axes, fully separating the values, and stacking. It is helpful if they also know about using averages such as medians or means to summarize group differences.

Activity Time: One class period

Objectives

- Understand that in comparing two groups using a numeric attribute, you can sometimes view one as an *additive shift* of the other.
- In comparing two distributions, recognize the additive constant as a “signal” that can be detected with indicators of center, such as clumps, hats, medians, or means.
- Recognize that some characteristics of distributions (such as the location of their centers) are evident in fairly small random samples, and the shifts between two distributions can be detected even in groups of unequal sizes.

Common Core Standards Addressed

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 the variation in estimates or predictions.

Grade 7, Statistics and Probability Standard 2

Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability.

Grade 7, Statistics and Probability Standard 3

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

- Ability to sketch and compare the shapes of different types of distributions (could be learned by doing the activity “Sketching Distributions”)

Materials

- Distribution Shift worksheet (one per student)
- **Post Office.tp**
- **Biased Scale.tp**
- **Pumpkin Farm.tp**

LESSON PLAN

Introduction (5 minutes)

Discuss how the cost of mailing letters is determined: Explain that the cost of postage is based on the weight of the letter. If you go to the post office, they put the letter on a scale, read out its weight in ounces, and look up the cost on a chart.

Open **Post Office.tp**. Have only the plot visible.

First-class letters often weigh between 1 and 13 ounces. Here is the kind of data we might get if we were to keep track of the weight of first-class letters. How would you describe the distribution of weights?

Explain that these weights were made in a TinkerPlots factory. Drag the lower-right corner of the sampler to make it visible. Click the **RUN** button to begin running it at a slow speed. Call the students' attention to the different widths of weight slices in the spinner. These correspond roughly to the shape of the distribution we get. You might also replace the spinner with a bars device, which shows the distribution of letter weights in a more traditional way. However, one advantage of using the spinner is that students have a better sense of how values are randomly selected, and can easily see that the thicker slices have higher probabilities of being landed on.

Change Repeat to 500 and run the sampler a few times to show the persistence of the shape and location of center. Ideally, students are now describing the center using concepts like the center clump, hat plot, median, or mean. If students are familiar with these concepts but don't know how to display them in TinkerPlots, this is a good time to show them. It can be informative to click the **Mean**, **Median**, and **Hat Plot** buttons in the upper plot toolbar and watch how these values change as more values are added to the sample. You can also click the **Divider** button to add a divider, and drag its borders to outline the center clump.

Modeling (15 minutes)

Display **Biased Scale.tp**.

Here's a TinkerPlots factory to model a post office with two scales for weighing letters, Scale A and Scale B. Let's watch what happens when people bring in letters to get weighed.

Run the sampler at slow speed so that students can come to understand how it works. It's critical that they see how the value for *DisplayWeight* is determined, by adding the value of *Bias* to the *ActualWeight* of the letter. Call their attention to these values in the results table, so students can see that the *DisplayWeight* is the sum of the *ActualWeight* and the *Bias*. Once students understand this, explain:

Post office workers weighed 500 letters, some on Scale A and some on Scale B. Here's the scale display weight of all the letters.

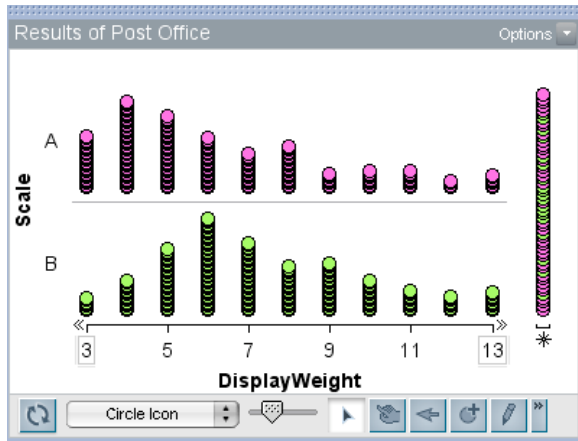
Increase the **Run Speed** slider's speed to quickly show the plot of all 500 letters' display weights. Make separate plots of the letters weighed on Scale A and Scale B by clicking the *Scale* attribute in the results table, then pulling any point up to separate the data into two separate plots. Make sure Scale A appears at the top of the plot, so that it is easy to see the relationship between the plot and the sampler.

Do you think we can tell the difference between the letters weighed on Scale A and those weighted on Scale B? How?

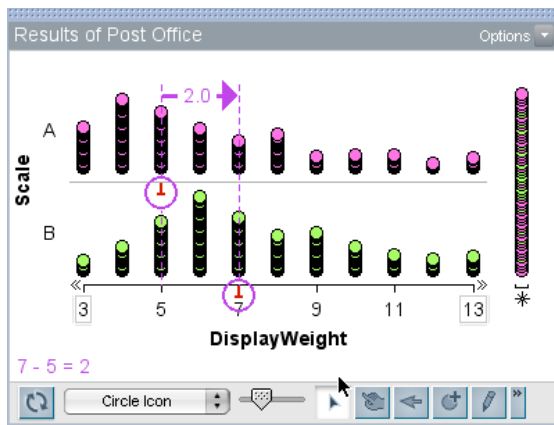
Encourage the description of one distribution being “shifted up” by one. Many students will pay most attention to the minimum and maximum of the distribution.

Do you think if you couldn't see the ends of the distribution, you could still spot the difference? How?

Hide the ends by double-clicking the values at the ends of the axes and setting the minimum to 3 and the maximum to 13. Ask students what tools they would use to spot the difference now.



In discussion with students, look for opportunities to use center indicators such as medians, means, hat plots, and dividers to help see the “shift of 2.” If students do not suggest it, introduce the median as a way to compare the two groups. You can use the **Ruler** tool to compute the difference between the two medians. It may be helpful to lock the ends of the ruler on the medians as shown here.



Drag *DisplayWeight* to the horizontal axis again so you can see the entire distribution. Run the sampler several times with the ruler locked on the medians to test how consistent the difference between the medians is. (The difference will almost always be 2.)

Suppose we were to make Scale B even more inaccurate, so that it adds 5 ounces to the actual weight of a letter. How would these two distributions compare then?

Replace 2 with 5 in Scale B's *Bias* spinner and repeat samplings. Again, use center indicators (medians) to see and describe the difference.

Suppose we were to reduce the number of letters weighed with the biased scale.

Drag the boundary between the two parts of the *Scale* spinner to make the likelihood of Scale B about $\frac{1}{3}$, and run the sampler again.

Are you still able to see the difference between the two groups even though the two groups are not the same size?

This is an important question, as many students will think you can't compare unequal groups. In this context, they should be able to see that they can still detect the difference.

Discuss and repeat running the sampler until students agree that even when the sample sizes are different, you can still detect the constant that is shifting weights measured on the biased scale upward.

Student Work at Computers (15 minutes)

Open the TinkerPlots document **Pumpkin Farm.tp**. Explain,

Here is a TinkerPlots factory to model a pumpkin farm where the farmer has decided to add a fertilizer to part of her crop.

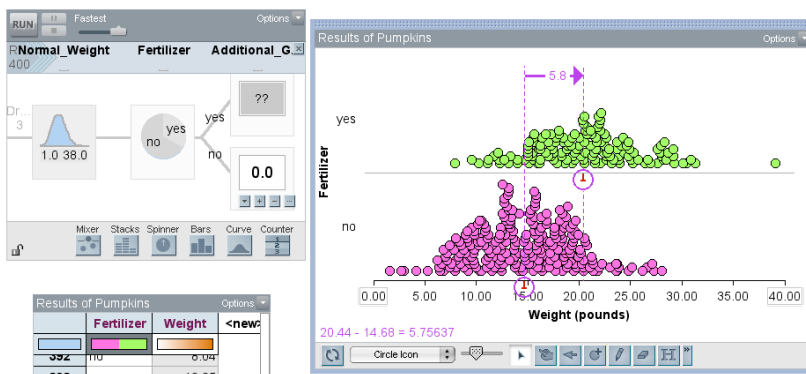
Show that the table lists the weight of each pumpkin and whether it had been fertilized. Tell students that they will decide whether the fertilizer made pumpkins weigh more, and if so, by how much.

Hand out the Distribution Shift worksheet and have students complete Steps 1-5.

Discussion (10 minutes)

Reveal to students that the pumpkins were made in a pumpkin factory by moving the Results of Pumpkins table and opening the sampler. Run the sampler at slow speed to show that it works much like the post office factory, but in this case the amount that is being added to the weight of the normal pumpkins is hidden.

How much additional growth do you think the fertilized pumpkins are getting?



Many students will offer estimates based on the differences between the groups they found in Step 4. Record a few of their responses on the board.

With the ruler locked on the medians, click the **RUN** button a few times so students can observe that for each sample they get a different result. Although the results vary, they do fall close together, indicating a difference of between 4 and 6 pounds. (To keep the graph from moving around due to slightly different ranges, double-click the axis end values and change them to 0 and 40.)

After students have observed the results from several samples, ask them to guess what number is in the box. Encourage them to give their estimate as a range rather than a specific value.

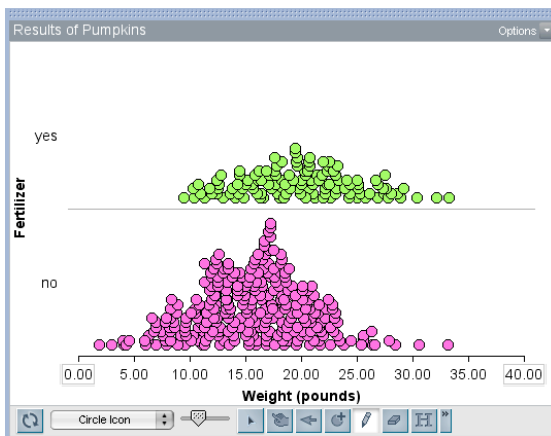
After students have given their estimates, show them the contents of the hidden box. To do this, hover over the upper *Additional_Growth* device and click the arrow in the lower-left corner to see the **Device Options** menu. Choose **Show Contents**.

Wrap-Up (5 minutes)

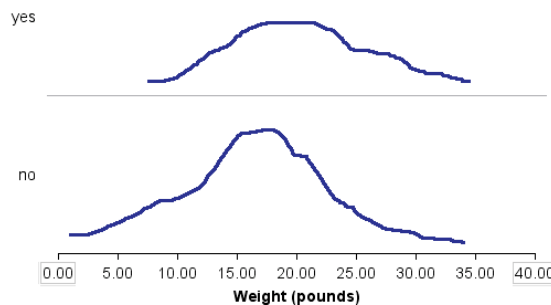
Have students explain why they think this activity is called “Distribution Shift.” What does it mean to *shift*?

ANSWERS

- Many students will make plots with the pumpkin weights in bins, making it more difficult to get a sense of the shape of the distribution. Encourage students to fully separate the values.



- Yes. Student sketches should show the distribution of fertilized pumpkins shifted slightly to the right of the distribution of unfertilized pumpkins. Their sketches should also indicate that there are fewer cases in the group of fertilized pumpkins. Students should avoid showing small details seen in their actual plots. Some students will also show the center clumps, clearly illustrating that the center clump of fertilized pumpkins is to the right of the center clump of the unfertilized pumpkins.





4. Answers will vary depending on the method of comparison students used. The median of the group of fertilized pumpkins is about 4 pounds more than the median of the group of unfertilized pumpkins, and the mean is about 4.3 pounds more. Students who compare center clumps may give answers between 2 and 7, depending on where they located the center clumps.
5. No, the relative group sizes do not affect the horizontal shift. Many students think you can't compare groups unless they are the same size. Students who understand that the size of the groups does not matter are likely developing a deeper perspective.