



## Lesson Overview

In this TI-Nspire lesson students are introduced to a histogram, the concept of bin width, and the notion that the frequency or number of data values in a bin is represented by the height of the corresponding bar. They explore how changing the bin width can change the story in the distribution of the data. Students compare different bin widths for the same data and to explore the effect of adding data points to a histogram.



Histograms are another useful tool for representing and analyzing data.

## Learning Goals

1. Describe the relationship between a dot plot and a histogram;
2. recognize a histogram is useful in representing large data sets;
3. identify the shape of a distribution of data represented in a histogram as skewed, symmetric, mound shaped, bimodal, or uniform, and identify characteristics of the distribution such as gaps, clusters, and outlying points;
4. approximate and interpret summary measures for data represented in a histogram.

## Prerequisite Knowledge

*Introduction to Histograms* is the seventh lesson in a series of lessons that investigates the statistical process. In this lesson, students investigate the effect changing the bin width has on the same set of data. Prior to working on this lesson students should have completed *Introduction to Data* and *Box Plots*. Students should understand:

- how to interpret data represented on a bar graph;
- how to interpret data represented on box plots and dot plots
- how to describe data represented on plots and graphs.

## Vocabulary

- **histogram:** a graphical display where the data is grouped into ranges and then plotted as bars
- **dot plot:** a graphical display of data using dots on a number line
- **symmetric:** when one side is the exact image or reflection of the other
- **skewed:** data that clusters towards one end of a graphical display
- **mound shaped:** data that clusters towards the middle of a graphical display
- **bimodal:** data distribution having two equal, most common values
- **uniform:** when the observations in a set of data are equally spread across the range of distribution

## Lesson Pacing

This lesson should take 50–90 minutes to complete with students, though you may choose to extend, as needed.

## Lesson Materials

- Compatible TI Technologies:



TI-Nspire CX Handhelds,



TI-Nspire Apps for iPad®,



TI-Nspire Software

- Introduction to Histograms\_Student.pdf
- Introduction to Histograms\_Student.doc
- Introduction to Histograms.tns
- Introduction to Histograms\_Teacher Notes
- To download the TI-Nspire activity (TNS file) and Student Activity sheet, go to <http://education.ti.com/go/buildingconcepts>.

## Class Instruction Key

The following question types are included throughout the lesson to assist you in guiding students in their exploration of the concept:



**Class Discussion:** Use these questions to help students communicate their understanding of the lesson. Encourage students to refer to the TNS activity as they explain their reasoning. Have students listen to your instructions. Look for student answers to reflect an understanding of the concept. Listen for opportunities to address understanding or misconceptions in student answers.



**Student Activity:** Have students break into small groups and work together to find answers to the student activity questions. Observe students as they work and guide them in addressing the learning goals of each lesson. Have students record their answers on their student activity sheet. Once students have finished, have groups discuss and/or present their findings. The student activity sheet can also be completed as a larger group activity, depending on the technology available in the classroom.



**Deeper Dive:** These questions are provided for additional student practice and to facilitate a deeper understanding and exploration of the content. Encourage students to explain what they are doing and to share their reasoning.



## Mathematical Background

In Lesson 1, *Introduction to Data*, students investigated dot plots of distributions of the maximum speeds of animals and their maximum life spans. They identified shapes, such as skewed, mound shaped, symmetric, bimodal, and also characteristics of some distributions such as gaps, clusters and outliers. They investigated box plots as a way to represent a five-number summary of the data in Lesson 3, *Box Plots*. Histograms, where sets of data values are grouped together into intervals, provide another way to represent and think about distributions of data and their shapes. Each interval has a vertical bar whose height is determined by the frequency or relative frequency of values contained in the interval. The width of these bars is up to the user and is called the bin width. A distinction between box plots and histograms is that the height of the bar represents the frequency of data values associated with that bin while in a box plot the height of the box does not represent any particular aspect of the data.

Histograms are particularly useful for very large data sets where a dot plot would be difficult to manage. The bin width can help the observer make sense of the story in the data or, in the case of too small or too large a bin width, obscure the story. If there are too many bars in the graph, a distribution can look noisy and cluttered. If the bin width is too large, patterns, gaps, and clusters may not be visible. No “magic” number of bars is correct; one of the advantages of dynamic interactive software is that the user can experiment with the bin width to see which best helps the observer understand the underlying shape of the data.

Common misconceptions related to histograms include confusing histograms and bar graphs; note that bar graphs have no inherent order and are used to indicate some feature of qualitative data while histograms are ordered according to a number line and are used with quantitative data. Related to the confusion between bar graphs and histograms, some think of histograms as displays of raw data with each bar standing for an individual observation. Some tend to judge variability by focusing on the varying heights of the bars, thinking of statistical variability in terms of frequencies, rather than data values.



## Part 1, Page 1.3

Focus: A histogram can be used to summarize and display a distribution of data.

Page 1.3 displays a dot plot of the number of pairs of shoes each student in a class of 23 owns.

**Select to add graph** displays a histogram with bin widths of 1.

The arrow keys on the screen or the keypad changes the bin widths.

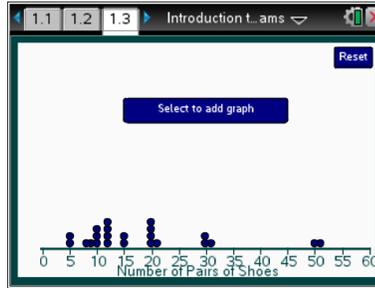
**Add Data** adds a new point at a specified value. Selecting any space under the axis will also add a point at that value.

Points can be dragged or moved using the arrow keys on the keypad.

**Deselect** deselects a highlighted point.

**Up/Down arrow** determines where tab and Left/Right arrows are active.

**Left/Right arrow:** change bin widths on the graph or moves the highlighted point on the bottom graph.



### TI-Nspire Technology Tips

**menu** accesses page options.

**tab** moves cycles through the bars on the top graph or the points on the bottom graph.

**enter** adds a point equal to the highlighted point.

**esc** releases the selected point.

**ctrl del** resets the page.



### Class Discussion

*The dot plot displays the number of pairs of shoes owned by the students in a sixth-grade class.*

- **What do you notice about the number of pairs of shoes the students owned?**
- **How many students were in the class? How did you find your answer?**
- **Remember the importance of thinking about shape, center and spread when talking about distributions of data. Describe the distribution on page 1.3.**

Answer: Some might be surprised that people own 50 pairs of shoes.

Answer: 23 students. I found the answer by counting the number of dots in the plot.

Answer: The distribution has two outliers that make it look kind of skewed right, but there is a big gap between owning 30 and 50 pairs of shoes. The number of pairs of shoes the students owned goes from 5 to 51 for a range of 46 pairs of shoes.

**Select to add graph.**

- **How does the new graph compare to the dot plot?**

Answer: The dots have been replaced by a vertical bar, and a vertical axis showing frequency has been added.



## Class Discussion (continued)

- **Select the bar at 30. How many dots does the bar represent, and what does this mean in terms of pairs of shoes?**  
Answer: The bar represents the two students who each own 30 pairs of shoes.
- **Explain how to interpret the interval represented by the bar. Why is it important to understand exactly which numbers are contained in the interval?**  
Answer:  $[30, 31)$  stands for the people who have at least 30 pairs of shoes but less than 31. It is important to know which edge point of the interval is in the bar and which is not; no point should belong to more than one bar.

**The new graph is called a histogram. Each bar represents the number of data values in an interval called the bin. The width of the interval is called the bin width. The height of the bar represents the number of data values in the bin.**

- **Select the right arrow under the number line in the middle once. How has the bar that contains 30 pairs of shoes changed?**  
Answer: The bar contains the dots for students with 30 and 31 pairs of shoes, and it has gotten taller, from a frequency of two to a frequency of three because all together, three people have 30 or 31 pairs of shoes.
- **Select the arrow again. Describe what happens each time you click the arrow.**  
Answer: The bin width changes.
- **What does the label frequency on the vertical axis represent?**  
Answer: The frequency tells you the number of people with that many pairs of shoes, which is really the height of the bar.

**Make the bin width as large as possible.**

- **How does this picture of the data compare to what you can see in the dot plot?**  
Answers will vary. You lose all of the gaps. You cannot tell if any students have between 31 and 40 pairs of shoes on the histogram. The histogram shows a skewed distribution without showing the peaks around 15 and 20.
- **What is the bin width for the first bar on the left?**  
Answer: The bin width goes from 0 up to 20.
- **How many people have 19 or fewer pairs of shoes? Describe two ways to find your answer.**  
Answer: 13 students have 19 or fewer pairs of shoes. You can highlight the bar and count the dots on the corresponding dot plot or you can read the frequency from the vertical axis.



## Class Discussion (continued)

**Make the bin widths 10.**

- **Which histogram, the one with bin width 10 or the one with bin width 20 gives a better picture of the actual distribution of the number of pairs of shoes students have? Explain your reasoning.**
- **How many people owned from 10 to 29 pairs of shoes? Explain how you can use the histogram to find your answer.**
- **Make the bin width 5. How does this compare to a bin width of 10?**
- **With bin width 5, move the point at 50 to 54. Describe how the histogram changes.**
- **Anita says the IQR should be from 10 to 20 caps. What would you say to Anita?**

Answer: The histogram with a bin width of 10 shows a gap between 40 and 50, which is not the whole gap but at least shows a gap. It also shows a cluster between 10 and 15 pairs of shoes, which you cannot see when the bin width is 20.

Answer: 14 people owned from 10 to 29 pairs of shoes. You can highlight the bars from 10 to 20 and from 20 to 29 and count the dots on the number line that would be in the bars or you can add the frequencies in the first two bars,  $5 + 9 = 14$ .

Answers will vary. When the bin width is 5, it is a better match for what is really happening with the data. You can see a larger gap before the last bar, a gap from 25 up to 30, and clumps from 10 up to 15 pairs of shoes and from 20 up to 25 pairs of shoes.

Answer: Nothing changes in the histogram.

Answer: She is giving the interval from the LQ to the UQ. The IQR is a number,  $UQ - LQ$ .

## Part 1, Page 2.2

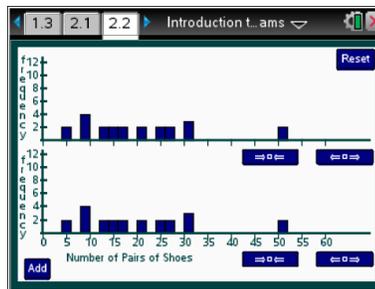
Focus: A histogram can be used to summarize and display a distribution of data.

Page 2.2 shows two histograms of the same data. The arrows on the screen or the arrow keys will change the bin widths on the corresponding graph.

**Add** allows you to type in a value, then **enter** to add a point to both plots.

**Left/Right** arrows change bin widths.

**Up/Down** arrows choose between graphs.



### TI-Nspire Technology Tips

**menu** accesses page options.

**tab** shows the frequency for the bins on the active graph.

**enter** opens add data and submits to graph.

**ctrl del** resets the page.



## Class Discussion

Page 2.2 displays identical histograms of the same data. Two students made comments about the histograms. What would you say to them? Give an example to help them see what you are talking about.

- **Kim noted that everybody but two people in the class had two pairs of shoes.**  
Answers may vary. Students should recognize that Kim is mixing up the variability in the number of pairs of shoes owned by all of the students with the frequency of how many students owned a certain number of pairs of shoes. Two people have twenty-five pairs of shoes, but two other people have only five pairs of shoes and two other people have as many as 50 pairs of shoes. So the range is from 5 to 50 pairs.
- **Sam said the class didn't have a lot of difference in the number of shoes they owned because the bins were all about the same height- between 2 and 4.**  
Answers may vary. Sam is making the same mistake as Kim, confusing the frequency of how many students owned a given number of pairs of shoes with the actual number of pairs of shoes students own.



## Student Activity Questions—Activity 1

### 1. Explain what a bin is and how to interpret the height of a bar associated with the bin.

Answer: The bin associated with a bar specifies an interval that is the base of the bar; all of the data values in the interval are included in the bar for that interval. The height of the bar indicates the number of data values in the bin. Students might exchange answers to this question to see if what the other person wrote makes sense and is clear.

### 2. Describe the difference in the shape of the two histograms if

#### a. the bin width on the top histogram is 7.

Answers may vary. The histogram on the top seems to be skewed right, while the one on the bottom except for the outlier is almost rectangular.

#### b. the bin width on the top histogram is as large as possible.

Answers may vary. The bin widths are 15, and again the histogram on top seems to be skewed right and you cannot see the gap between 30 and 31 pairs of shoes to 50 and 51 pairs of shoes. From what you can see, the three pairs of shoes owned by the students in that interval could be spread evenly across the interval.



## Student Activity Questions—Activity 1 (cont.)

- c. Add three students with 35 pairs of shoes and six students with 45 pairs of shoes. Select a bin width of 10. Describe the distribution.

Answers will vary. The distribution is almost rectangular, except for the last bar. That means that the same number of students, 6, are represented in each of the bars that have no difference in heights.

3. Move back to page 1.3 and add points for five students each having 40 pairs of shoes and for two students each having 35 pairs of shoes.

- a. Find a bin width that seems to give a good picture of the data.

Answer: A bin width of 5 seems to preserve the shape and represent the data fairly accurately.

- b. Explain why 40 is the tallest bar for bin widths of 1 but is not the tallest bar for bin widths of 5 and 10.

Answer: Because for bin widths of 5, the tallest bar is from 10 up to 15 and that bin contains the number of pairs of shoes for eight students, while the bar for 40 has the number of pairs of shoes for only five students. So the bar is taller from 10 up to 15 because it has more students.

- c. Describe how to use the histogram to determine the number of students who reported how many pairs of shoes they own.

Answer: Add up the frequency in each of the bars.

- d. The two students who had five pairs of shoes each bought 5 new pairs of shoes. Drag the points to update the distribution to account for the changes in the number of pairs of shoes for these students. Predict which bin will have the highest bar for bin width 5. For bin width 10? Explain your reasoning, and then check using the TNS activity.

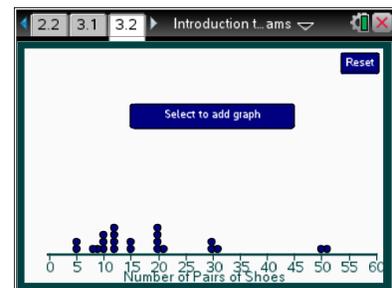
Answers may vary. The bars with bin width 10 up to 15 and bin width 10 up to 20 will be the tallest because those bars will represent the number of pairs of shoes for the most students.

## Part 2, Page 3.2

Focus: Measures of center and spread can be associated with histograms but only in a very general way.

On page 3.2, the buttons and menu options behave as they did on earlier pages.

**Mean $\pm$ MAD** and **Median, IQR** show the corresponding measures of center and spread for the data.





## Class Discussion

Have students...

**Go to page 3.2, create the histogram and make the bin width 5.**

- **Remember the mean of a set of data can be thought of as a balance point. Look at the histogram and estimate what you think the mean will be. Explain your reasoning.**
- **Allegra says she remembers that to find the mean you make all of the bars the same height. So she would move the dots around so the bars are the same height. What would you say to Allegra?**

- **Jorge said that to find the total number of pairs of shoes owned by students in class, you could add the numbers corresponding to the midpoint of each bar. Do you agree with Jorge? Why or why not?**
- **Use the dot plot to calculate the mean. (You may want to create a frequency table to organize your computations.) How did your estimate of the mean as balance point in the earlier question compare to the calculation?**

**Remember how to find the median for a set of data.**

- **Will the median number of pairs of shoes owned by students in class be smaller or larger than the mean? Explain your thinking.**

Look for/Listen for...

Answers may vary. Students might identify a value around 20 as the balance point for the distribution.

Answer: She is remembering what to do when you do not have the data on a number line- like the bags of dog food and the dogs. When you have data on the number line the dots have values, the number of pairs of shoes for a particular student. If two students each have three pairs of shoes and two students each have ten pairs of shoes, the mean is not two because there are two dots at each point (3 and 10); the mean is the total number of pairs of shoes divided by four students,  $3(2) + 10(2) = 26$ ;  
 $\frac{26}{4} = 6.5$  pairs of shoes per student for each of the four students.

Answers will vary. Jorge is not correct because he would only add up the 12 numbers from the number line that were the midpoints of the bars, but the bars represent 23 different students. He is ignoring the frequency.

Answer: Add up the number of pairs of shoes and divide by the number of students to find the mean number of pairs of shoes. The mean would be 18.6 or about 19 pairs of shoes per student, which is close to my estimate of 20 as the balance point.

Answer: The median will be smaller because the two students who owned 50 and 51 pairs of shoes will make the mean larger. You can see from the dot plot that more students have fewer than 19 pairs of shoes than have more than 19 pairs of shoes.



## Class Discussion (continued)

- Find the median number of pairs of shoes owned by students. Explain what the number represents and how you found it.**

Answer: You can tell by counting the dots in the dot plot there are 23 students all together. So the median is the 12<sup>th</sup> dot, which is 15 pairs of shoes per student; half of the students in the class had 15 or fewer pairs of shoes and half of the students in the class had 15 or more pairs of shoes.
- Can you find the median by looking at the histogram? Explain why or why not.**

Answer: You can estimate it by using the frequency in each bin, but you would only know which bin the median would be in and not the exact number for the median. If the bin width is 5, then the median would be in the third bin, from 15 up to 20, because there are 23 students all together and there are 11 students to the left of that bin and 10 students in the bins to the right of that bin. So the 12<sup>th</sup> student would be somewhere in the third bin.
- Remember the two measures for the spread or variability in a distribution of data, IQR and MAD. Which will be larger, the IQR or the interval determined by the mean $\pm$ MAD? Explain your reasoning.**

Answer: The interval determined by the mean  $\pm$  MAD will be larger because of the two students with 50 and 51 pairs of shoes. That makes the spread larger. The IQR does not count the value of the point, just the position so it is not affected by numbers far out at the end.
- Select Med, IQR. Describe how the median and IQR relate to the distribution of the number of pairs of shoes shown in the histogram. (Note the bar representing the length of the IQR can be dragged up to the histogram.)**

Answer: The distribution is slightly skewed right so the median is near the left of the distribution, in the third bar. The IQR is 11, and the interval associated with it includes the third, fourth and a part of the fifth bar, so it has to go from 10 up to 25.
- Select Mean $\pm$ MAD. What is the length of the interval determined by the mean $\pm$ MAD? How did you find it?**

Answer: The length of the interval is 18.6. You double the MAD.

## Class Discussion (continued)

- Explain any difference between the IQR and the length of the interval determined by the mean  $\pm$  the MAD for the number of pairs of shoes in the class.**

Answer: The length of the interval determined by the mean  $\pm$  MAD, 18.6, is larger than the IQR, 11, because of the two students who own 50 and 51 pairs of shoes. This makes the mean absolute deviation much larger because the deviations from the mean of 18.6 for the 50 and 51 are so large. But the IQR is the distance between the lower quartile and the upper quartile and is not affected by the values at 50 and 51, which are above the upper quartile.
- Set the bin width to 5. Add two data points that you think will make the IQR as long as the interval determined by the mean  $\pm$  MAD.**

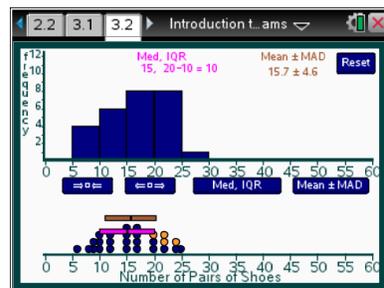
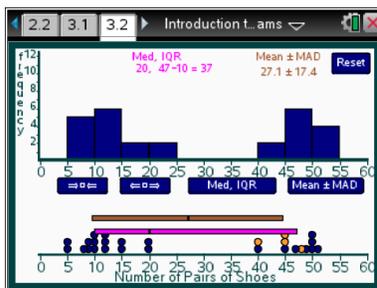
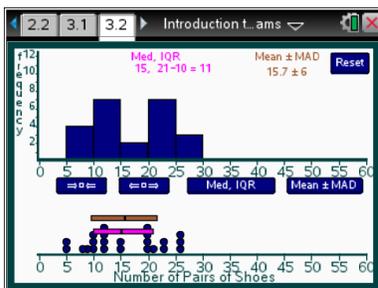
Answers will vary. Adding 35 and another 35 pairs of shoes extends the IQR to about as long as the interval determined by the mean  $\pm$  MAD.

## Student Activity Questions—Activity 2

- Work with a partner to create two reasonable distributions for the number of pairs of shoes owned by the students in a class, either by moving or adding points, 1) a distribution with little variability in the number of pairs of shoes owned by most of the class and 2) a distribution where there is a lot of variability in the number of pairs of shoes owned by the class.

Choose a bin width that seems best for your distribution. Describe your distribution (shape, center and spread). Explain why you think one of your distributions has very little variability and the other has a lot of variability.

Answers will vary. Have students share their distributions, either using Navigator to share with the whole class or by sharing screens on their handhelds or iPads with their group. Student reasons for having little variability should have to do with small measures for the IQR and mean  $\pm$  MAD indicating the spread is small. They may remember from Lesson 4, *Mean as Balance Point*, that the distributions with the largest deviations (spread) were those with the data at both ends of the distribution.





## Student Activity Questions—Activity 2 (continued)

2. Decide which of the following are always true, which are sometimes true, and which are never true. Give a reason for your answer.

a. **If the bin widths are greater than 1 in a histogram, it is not possible to compute the mean exactly.**

Answer: Always true because you do not know what the exactly values are in the bins so would have to estimate a value to represent those in each bin.

b. **The median will be in the tallest bar.**

Answer: Sometimes true, depending on how the data are ordered.

c. **You can determine the median from a histogram when the bin widths are greater than 1.**

Answer: Never true. You can find a bar that contains the median but you do not know the exact value as the median could be anywhere in the interval determined by the bin width.

d. **If all the bars are the same height, there is no variability in the data.**

Answer: The variability is determined by the spread of the data values, so it would be a function of the range of the values that make up the histogram, and of how the values were spread across the range.



## Deeper Dive

*Students might work with a partner on the question below.*

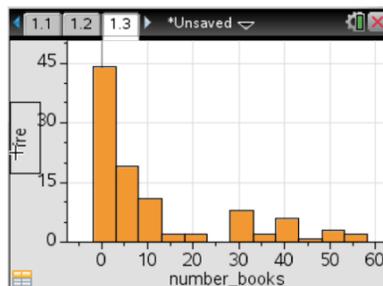
- *The following paragraph summarizes the findings of a 2014 survey related to the number of books people read. If 100 people were surveyed, use the information to create a possible histogram for the number of books people read in one year.*

**“One in 4 people aged 18 and older have not read a book in the last year. The mean number of books read or listened to in the past year is 12, and the median number is 4.5 books per year. 76% of American adults ages 18 and older said that they read at least one book in the past year, and 20% read at least 21 books per year.”**

Source: <http://www.pewinternet.org/2014/01/16/a-snapshot-of-reading-in-america-in-2013/>.

Answers will vary: The histograms should be skewed right because of the large difference in the mean and median, with about 18 people having read 0 to 1 books per year.

Sample answer:





## Deeper Dive – page 1.3

**Collect the number of pairs of shoes each person in your class has and enter it on page 1.3 by selecting a space below the number line and moving the point to the right place on the number line.**

- **How does your class data differ from the data for the class in the TNS activity?**

Answers will vary. Some may note that the yellow points representing the number of pairs of shoes owned by students in the class seems to be shifted to the left- overall, they own fewer pairs of shoes.

- **Find a bin width that seems reasonable for conveying the story about how many pairs of shoes students own. Explain why you think your choice is reasonable.**

Answers will vary. Student explanations should indicate that you can get a sense of the shape of the distribution as well as clusters and gaps.

- **Use your histogram to describe the number of pairs of shoes owned by both classes.**

Answers will vary. Students should use terms such as skewed, mound shaped, symmetric, outliers in their descriptions.



## Deeper Dive – page 2.2

**Find a method to estimate the mean of the data on page 2.2.**

Answers will vary. Students might choose to use the left end point of each interval, or the mean width (or midpoint) of each interval. The second method would give an estimate of 20.74 pairs of shoes as the mean number of pairs of shoes owned by students in the class.



## Deeper Dive – page 3.2

**On page 3.2, change the distribution by moving dots so the mean remains the same but the MAD is reduced by at least 3.**

Answers will vary. Students might choose to use the left end point of each interval, or the mean width (or midpoint) of each interval. The second method would give an estimate of 20.74 pairs of shoes as the mean number of pairs of shoes owned by students in the class.



## Sample Assessment Items

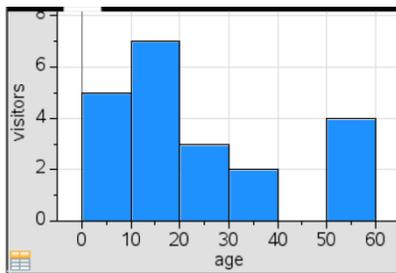
After completing the lesson, students should be able to answer the following types of questions. If students understand the concepts involved in the lesson, they should be able to answer the following questions without using the TNS activity.

1. This table shows the ages of 20 visitors at a library.

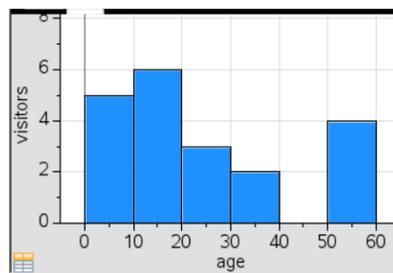
15	27	53	9	8
3	56	12	10	15
18	15	2	31	20
21	33	6	52	56

Which histogram shows the data?

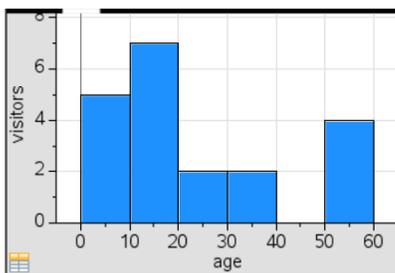
A.



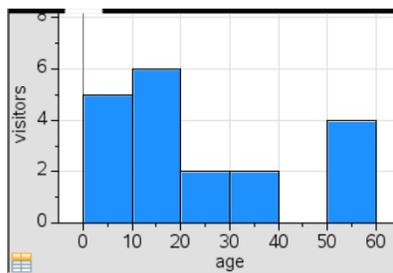
B.



C.



D.



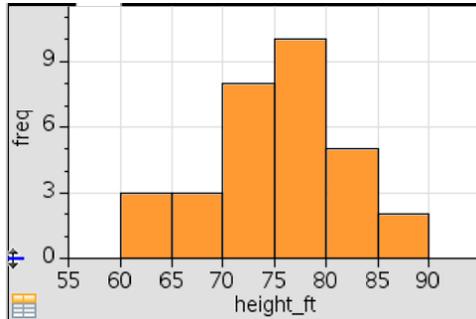
PARCC practice test 2014

**Answer: b**



2. The graph below displays the heights of cherry trees in a park in Washington DC. The first bin width is [60, 65).

**Height of Cherry Trees**



- a. How many cherry trees are in the park?

**Answer: 31**

- b. How many cherry trees are from 65 to 70 feet tall?

**Answer: 3**

- c. Which interval contains the median height of the cherry trees?

**Answer: [75, 80)**

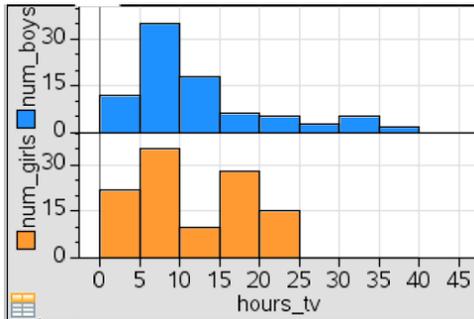
- d. Would you be surprised if the mean were 82 feet? Why or why not.

**Answer: It would be surprising to have a mean of 82 feet because only 7 of the 31 trees are taller than 80 feet, and the distribution is mound shaped and fairly symmetric so the mean and median would typically be close together.**



## Building Concepts: Introduction to Histograms TEACHER NOTES

3. The histograms display the number of hours of television per week watched by samples of 86 boys and 100 girls. Use the information in the histograms to identify each statement as true or false.



- a. The boys vary a lot more in the number of hours they watch television per week than the girls.

**Answer: True**

- b. About as many boys as girls watch television between 5 and 10 hours a week.

**Answer: True**

- c. Over half of the girls watch more than 15 hours of television a week.

**Answer: False**



## Student Activity Solutions

In these activities you will interpret and describe the distribution of data in a histogram. After completing the activities, discuss and/or present your findings to the rest of the class.



### Activity 1 [Page 2.2]

1. Explain what a bin is and how to interpret the height of a bar associated with the bin.

*Answer: The bin associated with a bar specifies an interval that is the base of the bar; all of the data values in the interval are included in the bar for that interval. The height of the bar indicates the number of data values in the bin. Students might exchange answers to this question to see if what the other person wrote makes sense and is clear.*

2. Describe the difference in the shape of the two histograms if

- a. the bin width on the top histogram is 7.

*Answers may vary. The histogram on the top seems to be skewed right, while the one on the bottom except for the outlier is almost rectangular.*

- b. the bin width on the top histogram is as large as possible.

*Answers may vary. The bin widths are 15, and again the histogram on top seems to be skewed right and you cannot see the gap between 30 and 31 pairs of shoes to 50 and 51 pairs of shoes. From what you can see, the three pairs of shoes owned by the students in that interval could be spread evenly across the interval.*

- c. Add three students with 35 pairs of shoes and six students with 45 pairs of shoes. Select a bin width of 10. Describe the distribution.

*Answers will vary. The distribution is almost rectangular, except for the last bar. That means that the same number of students, 6, are represented in each of the bars that have no difference in heights.*

3. Move back to page 1.3 and add points for five students each having 40 pairs of shoes and for two students each having 35 pairs of shoes.

- a. Find a bin width that seems to give a good picture of the data.

*Answer: A bin width of 5 seems to preserve the shape and represent the data fairly accurately.*

- b. Explain why 40 is the tallest bar for bin widths of 1 but is not the tallest bar for bin widths of 5 and 10.

*Answer: Because for bin widths of 5, the tallest bar is from 10 up to 15 and that bin contains the number of pairs of shoes for eight students, while the bar for 40 has the number of pairs of shoes for only five students. So the bar is taller from 10 up to 15 because it has more students.*

- c. Describe how to use the histogram to determine the number of students who reported how many pairs of shoes they own.

*Answer: Add up the frequency in each of the bars.*



# Building Concepts: Introduction to Histograms TEACHER NOTES

- d. The two students who had five pairs of shoes each bought 5 new pairs of shoes. Drag the points to update the distribution to account for the changes in the number of pairs of shoes for these students. Predict which bin will have the highest bar for bin width 5. For bin width 10? Explain your reasoning, and then check using the TNS activity.

*Answers may vary. The bars with bin width 10 up to 15 and bin width 10 up to 20 will be the tallest because those bars will represent the number of pairs of shoes for the most students.*

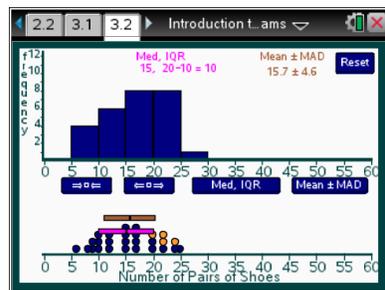
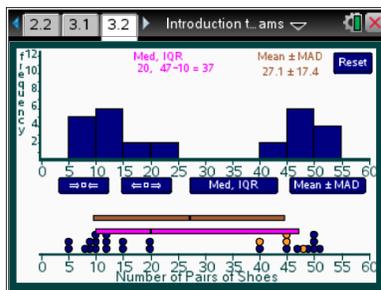
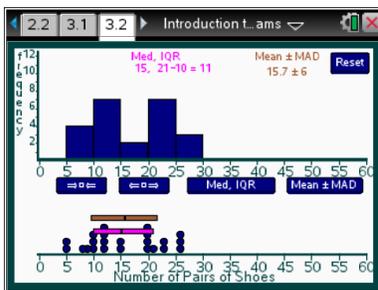


## Activity 2 [Page 3.2]

1. Work with a partner to create two reasonable distributions for the number of pairs of shoes owned by the students in a class, either by moving or adding points, 1) a distribution with little variability in the number of pairs of shoes owned by most of the class and 2) a distribution where there is a lot of variability in the number of pairs of shoes owned by the class.

Choose a bin width that seems best for your distribution. Describe your distribution (shape, center and spread). Explain why you think one of your distributions has very little variability and the other has a lot of variability.

*Answers will vary. Have students share their distributions, either using Navigator to share with the whole class or by sharing screens on their handhelds or iPads with their group. Student reasons for having little variability should have to do with small measures for the IQR and mean  $\pm$  MAD indicating the spread is small. They may remember from Lesson 4, Mean as Balance Point, that the distributions with the largest deviations (spread) were those with the data at both ends of the distribution.*



2. Decide which of the following are always true, which are sometimes true, and which are never true. Give a reason for your answer.

- a. If the bin widths are greater than 1 in a histogram, it is not possible to compute the mean exactly.

*Answer: Always true because you do not know what the exact values are in the bins so would have to estimate a value to represent those in each bin.*

- b. The median will be in the tallest bar.

*Answer: Sometimes true, depending on how the data are ordered.*

- c. You can determine the median from a histogram when the bin widths are greater than 1.

*Answer: Never true. You can find a bar that contains the median but you do not know the exact value as the median could be anywhere in the interval determined by the bin width.*



## Building Concepts: Introduction to Histograms TEACHER NOTES

- d. If all the bars are the same height, there is no variability in the data.

*Answer: The variability is determined by the spread of the data values, so it would be a function of the range of the values that make up the histogram, and of how the values were spread across the range.*