|  |
| --- |
|  |

Lesson 12: Describing the Center of a Distribution Using the Median

Student Outcomes

* Given a data set, students calculate the median of the data.
* Students estimate the percent of values above and below the median value.

Lesson Overview

The focus of this lesson is the median as a summary statistic to describe a data set. Students report the number of observations for both odd and even numbered sets of data. Informally, they consider the variability among three different data sets to assess a claim about typical behavior. In preparation for a later lesson on finding quartiles, students calculate the median of the values below the median and the median of the values above the median and estimate the approximate count/percent of values above and below the median. This lesson provides the background for the development of a box plot; however, this lesson is not about creating a box plot.

In this lesson students construct arguments and critique the reasoning of others. They respond to the reasoning of others in some of the tasks, distinguish correct reasoning from flawed reasoning, and explain why it is flawed. They also model with mathematics, apply mathematics to problems from everyday life, and interpret results in the context of the situation.

It should be noted that students should have access to calculators throughout this module.

Classwork

How do we summarize a data distribution? What provides us with a good description of the data? The following exercises help us to understand how a numerical summary answers these questions.

**Example 1 (2 minutes): The Median—A Typical Number**

The activity begins with a set of data displayed in a dot plot. Introduce the data presented in the example.

Example 1: The Median – A Typical Number

Suppose a chain restaurant (Restaurant A) advertises that a typical number of french fries in a large bag is . The graph shows the number of french fries in selected samples of large bags from Restaurant A.

****

**Number of French Fries in a Large Bag (Restaurant A)**

Using the data shown in the plot, students are asked to think about when it might be useful to separate a set of data into two parts that have the same number of elements. In other words, when would it be useful to know the point that separates the top half from the bottom half? The notion of median is developed by a set of questions and then defined.

Let students work independently on the exercises and confirm answers with a neighbor.

Exercises 1–3 (5 minutes)

Exercises 1–3

1. You just bought a large bag of fries from the restaurant. Do you think you have french fries? Why or why not?

The number seems to vary greatly from bag to bag. No bag even had fries, so mine probably will not. The restaurant sells french fries in bags that have from to per bag, so the claim that they typically have fries doesn’t seem right.

1. How many bags were in the sample?
2. Which of the following statements would seem to be true given the data? Explain your reasoning.
   1. Half of the bags had more than fries in them.
   2. Half of the bags had fewer than fries in them.
   3. More than half of the bags had more than fries in them.
   4. More than half of the bags had fewer than fries in them.
   5. If you got a random bag of fries, you could get as many as fries.

(a) and (b) are true because there are bags above fries and bags below fries. (e) is true because that happened once and so probably could happen again.

**Example 2 (3 minutes): The Median**

Read through the text with students.

Example 2: The Median

Sometimes it is useful to know what point separates a data distribution into two equal parts, where one part represents the larger “half” of the data values and the other part represents the smaller “half” of the data values. This point is called the median. When the data are arranged in order from smallest to largest, the same number of values will be above the median as are below the median.

As a class, work through the exercises one at a time. As students work through the problems, ask the following questions:

* How do you find the median if there are an even number of data points?
* About what fraction of the data values should be above the median? What fraction should be below the median?

Exercises 4–5 (5 Minutes)

Exercises 4–5

1. Suppose you were trying to convince your family that you needed a new pair of tennis shoes. After checking with your friends, you argued that half of them had more than four pairs of tennis shoes, and you only had two pairs. Give another example of when you might want to know that a data value is a half-way point? Explain your thinking.

Possible responses: When the data are about how much people earn, it would be interesting to know the amount that is less than what half of the people earn; if you are looking at the number of points earned in a competition, it would be good to know what number separates the top half of the competitors from the bottom.

1. Use the information from the dot plot in Example 1. The median number of fries was .
   1. What percent of the bags have more fries than the median? Less than the median?

percent or of the bags have more fries than the median, and or have fewer fries than the median.

* 1. Suppose the bag with fries was miscounted and there were only fries. Would the median change? Why or why not?

The median would not change because there would still be bags with fewer than fries and bags with more than fries.

* 1. Suppose the bag with fries really only had fries. Would the median change? Why or why not?

The median would change because now there would be bags that would have fewer than fries and only that have more than instead of the same number in both directions.

Exercises 6–7 (15 minutes): A Skewed Distribution

In this activity, students have to order the data before they find the median. There are values, so the median is the value with counts above and counts below. Another way to determine the median after ordering the data is to cross out the maximum and minimum values continuously until students reach one number in the middle if there are an odd number of data values, or two numbers for an even number of values. Students would then find the mean of the two values. The questions are designed to help students confront some common misconceptions: not ordering the data before counting to the middle; confusing median and mode (most frequent value); confusing median and midrange (half way between the maximum and the minimum). They also compute the mean and compare the median to the mean, noting that several bags with a low number of french fries pulled the mean down and so the median might be more reflective of the typical number of fries.

MP.3

Consider the following questions as students are completing this exercise:

* Why is it necessary to order the data before you find the median?
* Is the median connected to the range (maximum-minimum) of the data? Why or why not?
* What is the difference in the effect of very extreme values on the mean and on the median?

Exercises 6–7: A Skewed Distribution

1. The owner of the chain decided to check the number of french fries at another restaurant in the chain. Here is the data for Restaurant B: ,,,,,,, ,, , , , ,,,, , ,.
   1. How many bags of fries were counted?
   2. Sallee claims the median is as she sees that is the middle number in the data set listed above. She thinks half of the bags had fewer than fries. Do you think she would change her mind if the data were plotted in a dot plot? Why or why not?

You cannot find the median unless the data are ordered by size. Plotting the number of fries in each bag on a number line in a dot plot would order the data so you would probably get a different halfway point because the data above is not ordered from smallest to largest.

* 1. Jake said the median was . What would you say to Jake?

is the most common number of fries in the bags ( bags had fries), but it is not in the “middle” of the data, marking where the number of bags with fries more than and less than are the same.

* 1. Betse argued that the median was halfway between and or . Do you think she is right? Why or why not?

She is wrong because the median is not connected to the distance between points on the number line but are connected to finding a point that separates the data into two parts with the same number of values in each part.

* 1. Chris thought the median was . Do you agree? Why or why not?

Chris is correct because if you order the numbers, the middle number will be the number, with bags that have more than fries and bags with fewer than fries.

1. Calculate the mean and compare it to the median. What do you observe about the two values? If the mean and median are both measures of center, why do you think one of them is lower than the other?

The mean is and the median is . The bag with the fries lowered the value of the mean.

Exercises 8–10 (15 minutes): Finding Medians from Frequency Tables

**MP.4**

In this example, students find the median using a frequency table halfway between the and counts. They also find the medians of the top and bottom halves, the value from the top and from the bottom, as a precursor to finding an interquartile range in a later lesson. They will encounter repeated values in finding the quartiles. You may want students to write out the individual counts in a long ordered list. For example, the first counts would be as follows:

Median of the lower half

Then have students find the medians of each half by counting from the top and bottom of the list, noting that a value for bags with the same count can be in both halves. It might help to think about the individual bags – one of the bags with fries is in the first half, one of the bags with fries is in the second half, and one of the bags divides the two halves and marks the median of the data set. At this point, the important idea is that students get a sense of how to find a median: order the values and find a midpoint for the ordered values.

Exercises 8–10: Finding Medians from Frequency Tables

1. A third restaurant (Restaurant C) tallied a sample of bags of french fries and found the results below.

|  |  |
| --- | --- |
| Number of fries | Frequency |
|  | || |
|  | | |
|  | || |
|  | ||| |
|  | ||||| |
|  | |||| |
|  | | |
|  | | |
|  |  |
|  | ||| |
|  | ||| |
|  | | |

* 1. How many bags of fries did they count?
  2. What is the median number of fries for the sample of bags from this restaurant? Describe how you found your answer.

; I took half of , which was and then counted tallies from down to get to . I also counted up from to get to . The point halfway between and is the median.

1. Robere decided to divide the data into four parts. He found the median of the whole set.
   1. List the values of the bottom half. Find the median of these values.

Median of this half is .

* 1. List the values of the top half. Find the median of these values.

Median of this is .

1. Which of the three restaurants seems most likely to really have fries in a typical bag? Explain your thinking.

Answers will vary: Restaurant B seems to have most bags closest to because the middle half of the number of fries in bags covers a span of fries (from to ) with a median of ; Restaurant A goes from to fries for a span of with a median of , and Restaurant C goes from to for a span of but the median is . Some students might say Restaurant B because the median is .

Closing (3 minutes)

Lesson Summary

**In this lesson, you learned about a summary measure for a set of data called the *median*. To find a median you first have to order the data. The median is the midpoint of a set of ordered data; it separates the data into two parts with the same number of values below as above that point. For an even number of data values, you find the average of the two middle numbers; for an odd number of data values, you use the middle value. It is important to note that the median might not be a data value and that the median has nothing to do with a measure of distance. Medians are sometimes called a measure of the center of a frequency distribution but do not have to be the middle of the spread or range (maximum-minimum) of the data.**

Exit Ticket (5 minutes)

Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Lesson 12: Describing the Center of a Distribution Using the Median

Exit Ticket

1. What is the median age for the following data set representing the age of students requesting tickets for a summer band concert?
2. What is the median number of diseased trees from a data set of diseased trees on city blocks?
3. Describe how you would find the median for a set of data that has values. How would this be different if there were values?

Exit Ticket Sample Solutions

1. What is the median age for the following data set representing the age of students requesting tickets for a summer band concert?

The median is the value, or years old, as there are values less than and values greater than or equal to .

1. What is the median number of diseased trees from a data set of diseased trees on city blocks?

To find the median, the values first need to be ordered:

As there is an even number of data values, the median would be the mean of the and values: , or diseased trees.

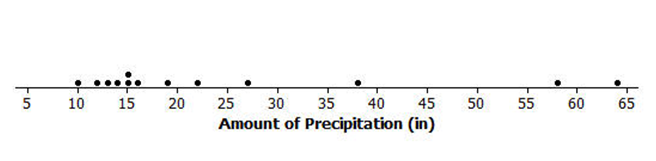
1. Describe how you would find the median for a set of data that has values. How would this be different if there were values?

Answers will vary; first you would order the data from smallest to largest. Because there are values, you would look for the value from the top or bottom. This would be the median with values above and values below. If the set had values, you would go halfway between the and values.

Problem Set Sample Solutions

1. The amount of precipitation in the western states in the U.S. is given in the table as well as the graph.

|  |  |
| --- | --- |
| State | Amount of Precipitation (in) |
| WA |  |
| OR |  |
| CA |  |
| MT |  |
| ID |  |
| WY |  |
| NV |  |
| UT |  |
| CO |  |
| AZ |  |
| NM |  |
| AK |  |
| HI |  |

 Data Source: http://www.currentresults.com/Weather/US/average-annual-state-precipitation.php

* 1. How do the amounts vary across the states?

Answers will vary: The spread is pretty large: inches. Nevada has the lowest at inches per year. Hawaii, Alaska, and Washington have more rain than most of the states; Hawaii has the most with inches followed by Alaska at inches.

* 1. Find the median. What does the median tell you about the amount of precipitation?

The median is inches; half of the states have more than inches of precipitation per year and half have less.

* 1. Use the median and the range to describe the average monthly precipitation in western states in the U.S.

The amount of precipitation varies from to inches per year. Half of the states have from to inches per year, but only two have more than inches.

* 1. Do you think the mean or median would be a better description of the typical amount of precipitation? Explain your thinking.

The mean at inches reflects the extreme values, while the median seems more typical at inches.

1. Identify the following as true or false. If a statement is false, give an example showing why.
   1. The median is always equal to one of the values in the data set.

False. If the numbers are and , the median is and it is not in the set.

* 1. The median is the midpoint between the smallest and largest values in the data set.

False. Look at the number of french fries per bag for Restaurant A above where the median is , which is not halfway between and ().

* 1. At most, half of the values in a data set have values less than the median.

True.

* 1. In a data set with different values, if you change the two smallest values of a data set to smaller values, the median will not be changed.

True.

* 1. If you add to every element of a data set, the median will not change.

False. The median will increase by as well. If the set is , , ,, , the median is ; for the set ,, ,, the median will be .

1. Make up a data set such that the following is true:
   1. The set has different values and the median is .

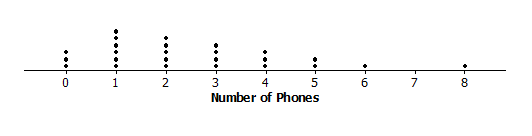
Answers will vary depending on the whether the numbers are whole numbers or fractions. If the numbers are whole numbers, the set would be , ,, , , , , , , , .

* 1. The set has values and the median is .

Answers will vary. One answer is to have all ’s.

* 1. The set has values and the median is the same as the smallest value.

Answers will vary. One answer is to have , ,, , , , .

1. The dot plot shows the number of landline phones that a sample of people have in their homes.
   1. How many people were in the sample?
   2. Why do you think three people have no landline phones in their homes?

Possible answers: Some people might only have cell phones, or some people may not be able to afford a phone.

* 1. Find the median number of phones for the people in the sample.

The median number of phones per home is .

* 1. Use the median and the range (maximum-minimum) to describe the distribution of the number of phones.

Possible answer: The median number of phones was per home, and over half of the people have fewer than phones in their homes. Three had none, and one house had phones.

1. The salaries of the Los Angeles Lakers for the 2012–2013 basketball season are given below.

|  |  |
| --- | --- |
| Player | Salary () |
| Kobe Bryant |  |
| Dwight Howard |  |
| Pau Gasol |  |
| Steve Nash |  |
| Metta World Peace |  |
| Steve Blake |  |
| Jordan Hill |  |
| Chris Duhon |  |
| Jodie Meeks |  |
| Earl Clark |  |
| Devin Ebanks |  |
| Darius Morris |  |
| Antawn Jamison |  |
| Robert Sacre |  |
| Darius Johnson-Odom |  |

Data Source: [www.basketball-reference.com/contracts/LAL.html](http://www.basketball-reference.com/contracts/LAL.html)

* 1. Just looking at the data, what do you notice about the salaries?

Possible answer: A few of the salaries for the big stars like Kobe are really big, while others are very small in comparison.

* 1. Find the median salary, and explain what it tells you about the salaries.

for Chris Duhon. Half of the players make more than and half make less than that.

* 1. Find the median of the lower half of the salaries and the median of the upper half of the salaries.

for the bottom half of the salaries; for the top half of the salaries.

* 1. Find the width of each of the following intervals. What do you notice about the size of the interval widths, and what does that tell you about the salaries?
     1. minimum salary to median of the lower half:
     2. median of the lower half to the median of the whole set:
     3. median of the whole set to the median of the upper half:
     4. median of the upper half to the highest salary:

The largest width is from the median of the upper half to the highest salary. The smaller salaries are closer together than the larger ones.

1. Use the salary table from above to answer the following.
   1. If you were to find the mean salary, how do you think it would compare to the median? Explain your reasoning.

Possible answer: The mean will be a lot larger than the median because when you add in the really big salaries, the size of the mean will increase a lot.

* 1. Which measure do you think would give a better picture of a typical salary for the Lakers, the mean or the median? Explain your thinking.

Possible answer: The median seems better as it is more typical of most of the salaries.