

NAME: _____

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Problem-Based Task 4.1.1: Summarizing Numerical Data Sets

Coaching

- a. What is the lowest price of the Algebra 1 textbook on SellUrBooks.com?
- b. What is the highest price of the Algebra 1 textbook on SellUrBooks?
- c. What is the mean price of the Algebra 1 textbook on SellUrBooks?
- d. What is the median value of the Algebra 1 textbook on SellUrBooks?
- e. Do you think the mean or median is a better measure of center for this data?
- f. What is the mean absolute deviation of the textbook prices on SellUrBooks?
- g. What is the interquartile range of the textbook prices on SellUrBooks?
- h. Do you think the mean absolute deviation or interquartile range is a better measure of variation?
- i. Each textbook seller only has one book to sell. When a student logs on to SellUrBooks, what price do you expect he or she will pay for the Algebra 1 textbook?
- j. Should Ella buy the book from SellUrBooks or iBuy.com? Use measures of center and variation to describe your answer.

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

Problem-Based Task 4.1.1: Summarizing Numerical Data Sets

Coaching Sample Responses

- a. What is the lowest price of the Algebra 1 textbook on SellUrBooks.com?

The lowest price is posted by seller 10: \$20.

- b. What is the highest price of the Algebra 1 textbook on SellUrBooks.com?

The highest price is posted by seller 3: \$56.

- c. What is the mean price of the Algebra 1 textbook on SellUrBooks.com?

The mean is the sum of all the prices divided by the total number of prices, 15.

$$\frac{32 + 44 + 56 + 29 + 48 + 45 + 34 + 44 + 42 + 20 + 46 + 43 + 40 + 52 + 41}{15} \approx \$41.07$$

- d. What is the median value of the Algebra 1 textbook on SellUrBooks.com?

First, order the prices from least to greatest.

Seller	Price
10	\$20.00
4	\$29.00
1	\$32.00
7	\$34.00
13	\$40.00
15	\$41.00
9	\$42.00
12	\$43.00
2	\$44.00
8	\$44.00
6	\$45.00
11	\$46.00
5	\$48.00
14	\$52.00
3	\$56.00

There are 15 data points, so the middle-most data point is the eighth price. The median value is \$43.00.

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

- e. Do you think the mean or median is a better measure of center for this data?

The mean is a good measure of center if there are no very high or very low values. In this data set, there is one low data value that may be reducing the mean. When comparing the mean and median, the mean is about \$2 less than the median. The median is not influenced by very high or very low values, and is a better measure of center for this data.

- f. What is the mean absolute deviation of the textbook prices on SellUrBooks?

First, find the absolute value of the difference between each price and the mean: $|\text{price} - \text{mean}|$.

$$| \$20.00 - \$41.07 | = \$21.07$$

$$| \$29.00 - \$41.07 | = \$12.07$$

$$| \$32.00 - \$41.07 | = \$9.07$$

$$| \$34.00 - \$41.07 | = \$7.07$$

$$| \$40.00 - \$41.07 | = \$1.07$$

$$| \$41.00 - \$41.07 | = \$0.07$$

$$| \$42.00 - \$41.07 | = \$0.93$$

$$| \$43.00 - \$41.07 | = \$1.93$$

$$| \$44.00 - \$41.07 | = \$2.93$$

$$| \$44.00 - \$41.07 | = \$2.93$$

$$| \$45.00 - \$41.07 | = \$3.93$$

$$| \$46.00 - \$41.07 | = \$4.93$$

$$| \$48.00 - \$41.07 | = \$6.93$$

$$| \$52.00 - \$41.07 | = \$10.93$$

$$| \$56.00 - \$41.07 | = \$14.93$$

Next, find the mean of these differences.

$$\frac{21.07 + 12.07 + 9.07 + \dots + 6.93 + 10.93 + 14.93}{15} = \frac{100.16}{15} \approx 6.72$$

The mean deviation is \$6.72.

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

- g. What is the interquartile range of the textbook prices on SellUrBooks?

The first quartile is the middle-most value of the lower half of the data. The data is divided by the median, the eighth data value of \$43. The first seven prices are the lower half of the data. The first quartile is in the middle of these prices; it is the fourth price, \$34. The third quartile is the middle-most value of the upper half of the data. The last seven prices are the upper half of the data. The third quartile is in the middle of these prices; it is the 12th price, \$46. The interquartile range is the difference between the third and first quartiles: $\$46 - \$34 = \$12$.

- h. Do you think the mean absolute deviation or interquartile range is a better measure of variation?

Since the median was chosen as the better measure of center, the interquartile range should be used as the measure of variation.

- i. Each textbook seller only has one book to sell. When a student logs on to SellUrBooks, what price do you expect he or she will pay for the Algebra 1 textbook?

The expected price is either the mean or median price. Since the median price was chosen as the better measure of center, a student can expect to pay \$43 on the SellUrBooks site.

- j. Should Ella buy the book from SellUrBooks or iBuy.com? Use measures of center or variation to describe your answer.

The expected price for a book on SellUrBooks is \$43, which is higher than the iBuy price of \$42. Ella will most likely pay more than \$42 if she purchases on the SellUrBooks site. She should buy from iBuy if she is worried that she won't be able to purchase the Algebra 1 textbook from one of the six sellers offering the book at a price less than \$42.

Recommended Closure Activity

Select one or more of the essential questions for a class discussion or as a journal entry prompt.

NAME: _____

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Practice 4.1.1: Summarizing Numerical Data Sets*

Elizabeth records her scores each time she goes bowling. The scores from her last 13 games are in the table below. Use the data to complete problems 1–6.

Game	Score
1	206
2	210
3	198
4	209
5	194
6	200
7	216
8	212
9	196
10	224
11	228
12	231
13	207

1. What is the mean of the data set?
2. What is the median of the data set?
3. Which measure of center best describes this data? Explain.
4. What is the mean absolute deviation of the data set?
5. What is the interquartile range of the data set?
6. Describe the variation of the data using both mean absolute deviation and interquartile range. Include comments about any striking deviations.

continued

NAME: _____

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Two science classes conducted an experiment. Each student measured the same amount of water in a beaker, and then heated the water until it boiled. The students recorded the time it took the water to boil in the tables below. Use the table to complete problems 7–10.

Class 1	
Student	Time in seconds
A	61
B	58
C	56
D	60
E	57
F	63
G	65
H	62

Class 2	
Student	Time in seconds
A	55
B	68
C	64
D	51
E	60
F	50
G	49
H	60
I	52

- For each class, what is the mean time it took the water to boil? For each class, what is the median time it took the water to boil?
- For each class, what is the interquartile range of the data? For each class, what is the mean absolute deviation?
- Why do you think the interquartile range is different for the two classes? Think about how the students set up the experiment and recorded the information.
- What overall patterns do you notice in the data sets? Are there any striking deviations?

NAME: _____

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Lesson 4.1.2: Representing Data Sets

Warm-Up 4.1.2

There are 10 ninth-grade classrooms. The number of students in each classroom is in the table below. Use the table to answer the questions about the number of students.

Classroom	Number of students
1	25
2	32
3	30
4	30
5	29
6	35
7	25
8	35
9	29
10	26

1. What is the median?
2. What is the first quartile?
3. What is the third quartile?
4. Which classroom has the least number of students?
5. Which classroom has the greatest number of students?
6. Are there any striking deviations in the data? Explain.

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

Lesson 4.1.2: Representing Data Sets

Common Core Georgia Performance Standard

MCC9–12.S.ID.1*

Warm-Up 4.1.2 Debrief

There are 10 ninth-grade classrooms. The number of students in each classroom is in the table below. Use the table to answer the questions about the number of students.

Classroom	Number of students
1	25
2	32
3	30
4	30
5	29
6	35
7	25
8	35
9	29
10	26

1. What is the median?

Order the data (the number of students in each classroom) from least to greatest.

Classroom	Number of students
1	25
7	25
10	26
5	29
9	29
3	30
4	30
2	32
6	35
8	35

The median is the middle-most value of the data set. If there is an odd number of data values, find the middle-most value. If there is an even number of data values, find the average of the

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

two middle-most values. The median number of students is the average of the fifth and sixth values, 29 and 30.

$$\frac{29 + 30}{2} = 29.5$$

2. What is the first quartile?

The median divides the data into two halves. The lower half of this data set is the first five data values: 25, 25, 26, 29, and 29. The first quartile is the middle-most value of the lower half of the data. The middle-most value of the lower half of the data is the third data value, 26.

3. What is the third quartile?

The upper half of this data set is the last five data values: 30, 30, 32, 35, and 35. The third quartile is the middle-most value of the upper half of the data. The middle-most value of the upper half of the data is the third value, 32.

4. Which classroom has the least number of students?

There are two classrooms with the least number of students. Both classrooms 1 and 7 have 25 students.

5. Which classroom has the greatest number of students?

There are two classrooms with the greatest number of students. Both classrooms 6 and 8 have 35 students.

6. Are there any striking deviations in the data? Explain.

Think about the typical number of students in a classroom. This can vary widely from school to school. In this data set though, the interquartile range is 6:

$$Q_3 - Q_1 = 32 - 26 = 6$$

Comparing the interquartile range and the median of 29.5 to the lowest data value, 25, the lowest data value is not far away from the mean in terms of the interquartile range. The lowest data value is not a striking deviation.

Now compare median and the interquartile range, 6. The highest data value, 35, is not far away from the mean in terms of the interquartile range. The highest data is not a striking deviation.

Connection to the Lesson

- In this lesson, students will plot data on the real number line using dot plots, histograms, and box plots.
- To create box plots from a given data set, students will need to determine the least data value, or minimum; the greatest data value, or maximum; the first quartile; the median; and the third quartile.

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

Prerequisite Skills

This lesson requires the use of the following skills:

- finding the median of a data set
- finding the first and third quartiles of a data set
- finding the minimum and maximum values of a data set
- plotting values on a real number line

Introduction

Data can be represented graphically using a number line. Graphs provide a visual representation of data; just by looking at a graph, you can quickly understand the spread and center of a data set. Dot plots and histograms show the frequency of a data value. In a **dot plot**, each data value is represented by a dot. The number of times a value is repeated corresponds to the number of dots above that value. In a **histogram**, the height of a rectangle above a value corresponds to the number of data values with that value. When looking at either a dot plot or histogram, it is easy to see both the most repeated data values and the spread of the data. If a data set is large, a histogram is easier to use because a single dot does not need to be drawn for each data value.

A **box plot** shows the minimum, maximum, first quartile, median, and third quartile of numerical data. The middle 50% of the data is represented with a box. Lines on either side of the box extend to the minimum and maximum data values. A box plot shows the range of data in a data set, and measures of center can be easily seen on a box plot. Box plots can be used to compare expected values of multiple data sets.

Key Concepts

- Numerical data can be represented graphically on the real number line.
- Dot plots and histograms show the frequency of each data value in a data set.
- Each data value in a data set is represented by a dot over that value in a dot plot.
- In a histogram, a rectangle is drawn above each value in a data set. The height of each rectangle corresponds to the number of data points with that value.
- A histogram can show the frequency of a range of values.
- The minimum, maximum, first quartile, median, and third quartile of a data set must be calculated before creating a box plot.

UNIT 4 • DESCRIBING DATA

Lesson 1: Working with a Single Measurement Variable

Instruction

- In a box plot, a rectangle is drawn starting at the first quartile and ending at the third quartile. The rectangle shows the middle 50% of the data set. The median is represented in the rectangle by a line. Whiskers are drawn from the rectangle to the minimum and maximum data values.
- A box plot shows more information about the expected value of a data set than a dot plot or histogram.
- A dot plot or histogram provides information about the size of a data set, which cannot be seen in a box plot.

Common Errors/Misconceptions

- confusing the mean and median
- using the mean in a box plot instead of the median
- incorrectly setting up the number line before creating a dot plot or histogram
- forgetting to include data values on a dot plot or histogram because they are not labeled on the number line