Building Concepts: Transforming Data

TEACHER NOTES

Lesson Overview

In this TI-Nspire lesson, students investigate what happens to the summary measures for the distribution of a set of data when a constant is added to each data element, or when each data element is multiplied by a non-zero constant.

Students discover what changes occur in the measures of center and spread, when a data set is transformed by shifting or scaling.

Learning Goals

1. Identify two common kinds of data transformations (shifting by a constant or scaling by a constant) involved in interpreting data;
2. recognize how these transformations affect measures of center and spread;
3. recognize that adding (subtracting) is an absolute change for each data value, while multiplying by a scale factor is a relative change for each data value.

Prerequisite Knowledge

Transforming Data is the tenth lesson in a series of lessons that investigates the statistical process. In this lesson, students investigate how data can be transformed and the effect such transformation has on the data set. This lesson builds on the concepts of the previous lessons. Prior to working on this lesson students should have completed *Analyzing Distributions*. Students should understand:

- how to read and interpret data on a dot plot;
- how to calculate the mean, interquartile range, and mean absolute deviation of a data set.

Vocabulary

- **transformation:** the changing of a data set through shifting or scaling
- **distribution:** describes the number of times each possible outcome appears in a sample or population.
- **mean:** the sum of all the data values in a set of data divided by the number of data values
- **median:** the value that separates the upper half of the distribution of a set of data values from the lower half
- **interquartile range (IQR):** the difference between the upper quartile and the lower quartile
- **mean absolute deviation (MAD):** the mean of the absolute values of all deviations from the mean of a set of data

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

- Transforming Data_Student.pdf
- Transforming Data _Student.doc
- Transforming Data.tns
- Transforming Data_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.

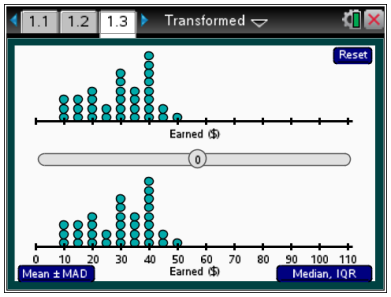
Mathematical Background

Transforming data is a fundamental process in data analysis. Often unwieldy data can be transformed by shifting or scaling to assist in its analysis. Such transformations typically occur when all the data is increased or decreased by a fixed number or the unit of measure for the data is changed. For example, dates are often transformed by subtracting to represent elapsed time: years since 1950 rather than the raw date. Also, when the units of measure of the heights of students in a seventh grade class are changed from inches to centimeters, the mean and median are multiplied by the scale factor 2.54 cm/inch.

When data are transformed by adding a constant, the change in the data is an absolute change; the effect on the magnitude of every data value is the same. When data are transformed by multiplying by a common factor, the change is relative; the change in the magnitude of the data varies depending on the data. For example, if the minimum value is 1 and the maximum value is 100, adding a constant such as 5 to each value will produce 6 and 105, an increase of 5 for both. However, multiplying the two values by a scale factor of $\frac{1}{2}$ will reduce the 1 to $\frac{1}{2}$, a decrease of $\frac{1}{2}$ a unit, and will reduce the 100 to 50, a decrease of 50 units.

Part 1, Page 1.3

Focus: Adding (subtracting) a common value to each element of a data set will shift the center by that value but will not change the measure of spread.



On page 1.3, the data displayed in the dot plots are the hourly wages for employees of a computer game company.

Mean+/-MAD and **Median, IQR** display the corresponding measures on both plots.

Dragging the point on the bar or using the left/right arrow keys on the handheld will change the employee hourly wage data on the upper dot plot. Points in the data set can be moved by dragging, or using tab and the left/right arrow keys.

Arrows chooses whether the Left/Right arrow keys control the slider or the data values.

TI-Nspire Technology Tips

menu accesses page options.

tab moves among the buttons or among the data values.

Up/Down arrows: control tab actions.

ctrl del resets the page.

Class Discussion

On page 1.3, the distribution represents the number of employees at each hourly wage in a computer game company.

- ***What are the minimum and maximum hourly wages for the employees in the company? How many employees have these hourly wages?***

  Answer: Three employees have the minimum wage of $10 per hour. One employee has the maximum wage of $50 per hour.

- ***Remember that the mean is the balance point of a distribution. Make a conjecture about the mean hourly wage for employees in this company. Use* Mean+/-MAD *button to check your conjecture.***

  Answer may vary. The mean hourly wage is $29.80 per hour.

- ***What do you think will happen to the distribution if the salary for every employee is increased by $5? Decreased by $5? Use the slider to check your conjecture.***

  Answer: The entire distribution moves to the right $5 if the salary for every employee is increased by $5. If the salaries are decreased by $5, the distribution moves to the left by $5.

Class Discussion (continued)

- ***How does the median change when you increase or decrease the hourly wage of all the employees by $5? Explain why this happens.***

  Answer: The median is shifted $5 to the right when all hourly wages are increased by $5 and $5 to the left when all hourly wages are decreased by $5. The median is the middle value of the data, and when you shift the each element of the data by the same amount, the order of the items does not change. The value that is in the middle stays in the middle and is just shifted by that same amount.

- ***Tiffany claims that the IQR increases by $5 if each employee's hourly wage is increased by $5. Is she correct? Why or why not?***

  Answer: Tiffany is incorrect. The IQR does not change. All the points move the same amount, so the median and the LQ and UQ will move $5 leaving the interval the same length. So the IQR is not changed since the distance between LQ and UQ remains the same.

- ***Do the mean and MAD change? If so, describe how.***

  Answer: The mean shifts to the right by $5, but since all the values change the same amount, the MAD does not change.

Student Activity Questions—Activity 1

1. On page 1.3, suppose the manager decides to raise the minimum wages to $15 an hour.

a. How many employees will be affected by the raise?

Answer: three, who were earning $10 an hour.

b. What will happen to the range of the distribution?

Answer: The range will be $35, from $15 to $50, instead of $40, from $10 to $50.

c. Describe what will happen to the median salary and the IQR. Explain your reasoning.

Answer. The median and IQR will not change because moving three data values to 15 does not change the LQ and so will not affect the median or the IQR.

d. Describe what will happen to the mean wage and the typical spread from the mean. Explain your reasoning.

Answer: The mean will increase because you are adding $15 to the overall sum of the employees' wages, so the mean will be a little less than $0.50 higher $\left(\frac{\text{change in total dollars}}{\text{number of employees}} = \frac{15}{33}\right)$. The typical spread from the mean will decrease because you have reduced the magnitude of the deviations from the mean when the data are moved closer together.

Student Activity Questions—Activity 1 (continued)

2. Identify each of the following as true or false. Give a reason for your answer.

If a value is added to each element in a data set,

a. the spread of the data will not change.

b. the shape of the data will be different.

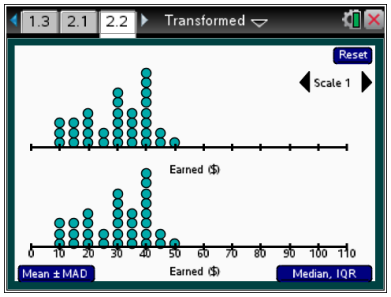
c. the center of the data distribution will change.

d. the maximum value will be increased by the amount added.

Answer: a, c, and d are true because each data value in the distribution has been shifted to the right or left by the same amount, so the center will change, but the measures of spread will stay the same. b is false because each value has been changed by exactly the same amount so the shape will be the same.

Part 2, Page 2.2

Focus: Multiplying the values in a data set will change the values by different absolute amounts and will change both measures of center and spread.



On page 2.2, selecting **Scale** will rescale all of the hourly wages by a scale factor from $\frac{1}{5}$ to 2.

The menu and buttons function in the same way as those on page 1.3.

Class Discussion

The dot plots on page 2.2 represent the wages of the employees of the same company. The manager gives overtime wages that are 1.5 times the regular wages.

- ***How much will the employees who make the least earn per hour in overtime?***

  Answer: The employees who make the least, $10 an hour, will make $15 an hour in overtime.

- ***Miranda said that everyone in the company will get the same increase in wages when they are paid overtime. What would you say to Miranda?***

  Answers may vary. She is right if she is thinking that everyone will get their salary plus half as much of their salary in overtime wages, but she is wrong in terms of how much money they will earn. The ones who earn the least will only make $15 an hour, a $5 increase, while the ones who earn the most will make $75 an hour, a $25 increase.

Class Discussion (continued)

Have students…	Look for/Listen for…
• ***Make a conjecture about what will happen to the distribution when the overtime hourly wages are computed. Change the scale to 1.5 to check your conjecture.***	Answer: Conjectures may vary. A typical conjecture would be to think the distribution moves to the right 1.5 units. The dot plot does move to the right, but the dots are more spread out.
Make a conjecture to answer of the following. Then use the TNS activity to check your conjecture.	
• ***What is the mean overtime wage for employees in the company? The typical difference from the mean (the MAD)***	Answers will vary. The mean and the MAD will both be approximately 1.5 times the original mean and the MAD, from a mean wage of \$29.80 to a mean of \$44.70 and from a MAD of \$4.60 to \$6.90.
• ***How do the median and the IQR for the distribution of overtime wages compare to the median and the IQR for the distribution of regular wages?***	Answer: The median and the IQR are multiplied by 1.5 and, therefore, change by that amount, from a median wage of \$30 to \$45 and the IQR from \$20 to \$30.
• ***How would you describe the typical wages of employees in the company before any overtime? After the overtime?***	Answers will vary. Some may want to use a measure of center, and for these data, they are about the same, around \$30. Others may want to use an interval such as from about \$20 to \$40 (using the IQR) or from about \$25 to \$35 (using the MAD). With the overtime, the typical wage is about \$45,or if they are using an interval, from about \$30 to \$60 (using IQR) or from about \$38 to \$52 (using MAD)
Suppose the distribution on page 2.2 represented the height in feet of the dinosaurs in a movie. A company is making a model of some of the dinosaurs to put into a theme park and want the mean height to be reasonable for children to climb on.	
• ***The company is trying to decide what scale to use for the dinosaurs in the park. Use the TNS activity to help you decide what scale you would recommend.***	Answers will vary. Using a scale of $\frac{1}{5}$ would bring the mean height of the dinosaurs to about six feet and the maximum height of 10 feet, which would be reasonable for older children to climb on. This would mean that some of the dinosaurs would be only two feet tall, which would be ok for smaller children.

Class Discussion (continued)

- ***Describe the height of a typical toy dinosaur using your scale.***

  Answers will vary depending on the scale chosen. For a scale of $\frac{1}{5}$, the typical height would be about 6 feet, and half of the dinosaurs would be between 4 and 8 feet tall.

Student Activity Questions—Activity 2

1. The holiday hourly wage is double the regular hourly wage.

a. Describe what happens to the dot plot when the holiday hourly wages are computed.

Answer: The dots spread out and move to the right.

b. How do the measures of center change for holiday wages?

Answer: The mean and median wages are doubled, to about \$60.

c. Which employees earned the most money during the holiday season compared to what they usually earned?

Answer: The employees who earned the most before have a greater increase than those who had a small wage (The minimum wage went to \$20 and the maximum to \$100). In terms of percent change, the changes would all be the same, 100%; percent change is not the same as the absolute total change.

2. The state tax for employees of this company is 20%. What is the mean amount of tax employees pay on their standard hourly wage?

Answer: The mean will be $\frac{1}{5}$ of the original mean, $\frac{1}{5} \times \$29.80 = \$5.96$.

3. How would you transform the prices of a set of items in a store so that each of the following is true:

a. The middle half of the prices would double in cost.

Answer: Double each of the prices.

b. The mean price would be \$5 higher.

Answer: Add 5 to each price.

c. The typical deviation from the mean price would be one third as much as in the original set.

Answer: Take one third of all the prices.

d. The median price would remain the same.

Answer: This is not possible unless you add 0 or multiply by 1.

Deeper Dive – pages 1.3 and 2.2

Which of the following would increase the total hourly wages of the computer game company employees the most? Explain your reasoning.

a. The hourly wage for each employee is increased by $20.

b. The hourly wage for each employee is multiplied by 1.5.

Answer: For a, when the hourly wage is increased by $20, the mean would increase by $20 to $\$29.80 + \$20 = \$49.80$.

For b, when the hourly wage is multiplied by 1.5, then the mean would increase to $\$29.80 \times 1.5 = \$44.70$

Since the mean is fair share value, adding $20 to each employee's hourly wage would yield a total hourly wage of $\$49.80 \times 33 = \$163.40$ while multiplying each employee's hourly wage by 1.5 would yield $\$44.70 \times 33 = \$1475.10$. So adding $20 would give the greatest increase in total hourly wage.

Deeper Dive – page 2.2

How can you transform a set of data so the IQR is doubled, but the mean remains the same?

Answer: Double each value and shift the data to the left the difference between the original mean and the doubled mean.

Deeper Dive

Consider the hourly wages of the employees of the computer game company. The manager laid off the three people who earned the most.

- ***Find the new mean and median hourly wages. Explain your reasoning.***

  Answer. The original mean was $29.80 for 33 employees. So the total amount in hourly wages would be $33(\$29.80) = \$983.40$. The three highest employees earned a total of $140 so the new total amount would be $843.40 and dividing this among the 30 remaining employees would give a new mean wage of about $28.11. The median hourly wage would remain the same, about $30, because half way between the wages of the 15^{th} and 16^{th} employee is at $30.

- ***If the manager gave every remaining employee a $4 raise, would he be spending more or less money than before he laid off the three people? Justify your answer.***

  Answer: He saved $140 in the layoffs; if he gave 30 people a $4 raise, he would be spending $120 so would save money.

- ***Suppose instead of giving the 30 remaining employees a $4 raise, he gave them a 20% increase. Would he be spending more or less than before he laid off the three people?***

  Answer: A 20% increase would increase the total he paid the 30 employees to $1012.08, which would result in a greater raise than if he gave a $4 raise.

Deeper Dive (continued)

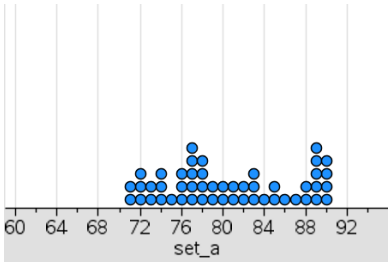
- ***Suppose instead of either of those plans, he divided the savings among the next highest paid of the 30 remaining employees. Would the new mean hourly wages return to the original mean? Why or why not?***

  Answer. It would not be the same as the original. The total amount of money in wages would be the same, but this time it is shared among 30 people not 33 so the mean would be slightly larger than the original mean.

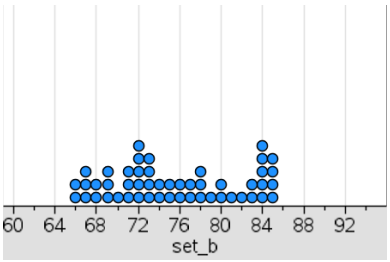
Sample Assessment Items

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. Consider the two dot plots below.



Set A



Set B

   a. How was the distribution in Set A changed to create Set B?

      Answer: Each value in Set A was decreased by 5.

   b. If the median of Set A was 79.5 and the interquartile range was 10, what is the median and interquartile range of Set B?

      Answer: median = 74.5; IQR = 10

2. A store owner decided to cut all of his prices by 1/3 for a big sale. If the mean cost of an item in the store was $24 before the price cut, what was the mean price after the cut?

   a. $36

   b. $16

   c. $12

   d. $8

   Answer: b. $16

3. One of the problems on a mathematics test was incorrectly stated. The teacher decided to add 3 bonus points to everyone's test score. If the mean grade on the test was originally 84 with a mean absolute deviation (MAD) of 10, what would the new mean and MAD be?

   Answer: The mean would be 87 and the MAD would be 10.

4. The Real Burger Company has 5 full time employees. The hourly wages of the employees are $12, $18, $19, $21, and $50.

   a. What is the mean salary for the 5 employees?

      Answer: $24

   b. Which value seems most reasonable for the mean absolute deviation of the salaries from the mean: $2, $10, $50?

      Answer: $10

   c. What is the median salary for the 5 employees?

      Answer: $19

   d. If the restaurant gave every employee a $2 raise, which would change the most, the median salary or the mean salary?

      Answer: They would both change by an increase of $2

5. Alex got a larger raise than Sylvia, but she took more money home. Explain how this might happen. (Use an example to support your thinking.)

   Answer: Alex might have earned a small amount like $100 a week and gotten a 5% raise, which would be $5. Sylvia might have earned more, like $200 and gotten a smaller percent raise like 2% so she would have gotten a $4 raise but would have brought home $204.

Student Activity Solutions

In these activities you will explain how data transformations affect measures of center and spread. After completing the activities, discuss and/or present your findings to the rest of the class.

Activity 1 [Page 1.3]

1. On page 1.3, suppose the manager decides to raise the minimum wages to $15 an hour.

   a. How many employees will be affected by the raise?

      Answer: three, who were earning $10 an hour.

   b. What will happen to the range of the distribution?

      Answer: The range will be $35, from $15 to $50, instead of $40, from $10 to $50.

   c. Describe what will happen to the median salary and the IQR. Explain your reasoning.

      Answer. The median and IQR will not change because moving three data values to 15 does not change the LQ and so will not affect the median or the IQR.

   d. Describe what will happen to the mean wage and the typical spread from the mean. Explain your reasoning.

      Answer: The mean will increase because you are adding $15 to the overall sum of the employees' wages, so the mean will be a little less than $0.50 higher $\left(\frac{\text{change in total dollars}}{\text{number of employees}} = \frac{15}{33}\right)$*. The typical spread from the mean will decrease because you have reduced the magnitude of the deviations from the mean when the data are moved closer together.*

2. Identify each of the following as true or false. Give a reason for your answer.

   If a value is added to each element in a data set,

   a. the spread of the data will not change.

   b. the shape of the data will be different.

   c. the center of the data distribution will change.

   d. the maximum value will be increased by the amount added.

   Answer: a, c, and d are true because each data value in the distribution has been shifted to the right or left by the same amount, so the center will change, but the measures of spread will stay the same. b) is false because each value has been changed by exactly the same amount so the shape will be the same.

Activity 2 [Page 2.2]

1. The holiday hourly wage is double the regular hourly wage.

   a. Describe what happens to the dot plot when the holiday hourly wages are computed.

      Answer: The dots spread out and move to the right.

   b. How do the measures of center change for holiday wages?

      Answer: The mean and median wages are doubled, to about $60.

   c. Which employees earned the most money during the holiday season compared to what they usually earned?

      Answer: The employees who earned the most before have a greater increase than those who had a small wage (The minimum wage went to $20 and the maximum to $100). In terms of percent change, the changes would all be the same, 100%; percent change is not the same as the absolute total change.

2. The state tax for employees of this company is 20%. What is the mean amount of tax employees pay on their standard hourly wage?

   Answer: The mean will be $\frac{1}{5}$ *of the original mean* $\frac{1}{5} \times \$29.80 = \$5.96$.

3. How would you transform the prices of a set of items in a store so that each of the following is true:

   a. The middle half of the prices would double in cost.

      Answer: Double each of the prices.

   b. The mean price would be $5 higher.

      Answer: Add 5 to each price.

   c. The typical deviation from the mean price would be one third as much as in the original set.

      Answer: Take one third of all the prices.

   d. The median price would remain the same.

      Answer: This is not possible unless you add 0 or multiply by 1.