Errata - April 2013

Bridging the Gap Between Common Core State Standards and Teaching Statistics

Page xvi

Linking Grade Levels and the Common Core State Standards

Investigation	Grades K–5	Grade 6	Grade 7	Grade 8
1.1		6.SP.1		
2.1	K.MD.3, 1.MD.4, 2.MD.10			
2.2	K.MD.3, 1.MD.4, 2.MD.10			
2.3	K.MD.3, 1.MD.4, 2.MD.10			
2.4	K.MD.3, 1.MD.4, 2.MD.10			
3.1	K.CC.7			
3.2		6.SP.1-4		
3.3		6.EE.2, 6.SP.1-4		
3.4		6.SP.1-5		
4.1		6.SP.1-5		
4.2		6.SP.1-5	7.SP.3	
4.3		6.SP.1 and 5, 6.RP.1		8.SP.1
4.4		6.RP.3c, 6.SP.3		8.SP.4
5.1				8.SP.1
5.2				8.SP.1
5.3				8.SP.1 and 2
5.4				8.F.3 and 4, 8.SP.2 and 3
6.1			7.SP.5	
6.2		6.SP.1 and 2	7.SP.5 and 8a	
6.3			7.SP.5, 7.SP.7b, 7.SP.8	
6.4			7.SP.6	

Common Core State Standards Grade Level Content

6.EE.2 Write, read, and evaluate expressions in which letters stand for numbers.

6.SP.1 Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers.

6.SP.2 Understand that a set of data collected to answer a statistical question has a distribution that can be described by its center, spread, and overall shape.

6.SP.3 Recognize that a measure of center for a numerical data set summarizes all its values with a single number, while a measure of variation describes how its values vary with a single number.

6.SP.4 Display numerical data in plots on a number line, including dotplots, histograms, and boxplots.

NCTM Principles and Standards for School Mathematics

Data Analysis and Probability

Grades 3-5 All students should describe the shape and important features of a set of data and compare related data sets, with an emphasis on how the data are distributed.

Grades 6-8 All students should formulate questions, design studies, and collect data about a characteristic shared by two populations or different characteristics within one population; find, use, and interpret measures of center and spread, including mean and interquartile range.

Materials

- Sticky notes
- Interlocking cubes (e.g., Unifix cubes)
- Chart paper

Estimated Time

Two days

Sum of the absolute deviations = 8

Mean absolute deviation = 8/10 = .8

On average, the shoe lengths are 0.8 cm from the mean length of 18 cm.

Extensions

1. An alternative approach to developing the mean as a balance point is to have your students look at the dotplot and make a guess as to where a fulcrum should be placed to balance the graph. Suppose their guess for the balance point is 18. The first five rows of a table that shows the deviations of the data from their guess of 18 would be the following:

Table for Calculating Deviations

Shoe Length	Shoe Length - 18
20	20 - 18 = 2
17	17 - 18 = -1
17	17 - 18 = -1
19	19 - 18 = 1
18	18 - 18 = 0

Using 18 as a guess for the mean, the sum of the deviations of all 25 shoe lengths is -7. A properly placed fulcrum would be one such that the total of the negative deviations balances (cancels out) the total of the positive deviations. Since the sum was -7, the guess for the balance point was too high.

Have the students try 17. The sum of the deviations using 17 as the guess for the mean is 18, which would mean the guess of 17 was too low.

Trying 17.7 produces a total of 0.5. This would mean that 17.7 would be very close to the actual balance point. We know the mean is 17.72, so that if we now tried 17.72, the sum of all 25 deviations would turn out to be 0.

2. Add famous athletes' shoe sizes to your class data. For example, Shaquille O'Neal has shoe size 22 (about 47 cm); Michael Phelps has 14 (about 30.5 cm); and Lebron James is 16 (about 31.8 cm). How are the original conclusions affected by the inclusion of these three data points?

3. Investigate any relationship between shoe length in centimeters and shoe size. From your findings, what would a student's shoe size be if the student has a shoe length of 17 cm?

discuss gaps and the reason the spread in the third group is so wide. Note that it is due to 34.3 being so much higher than the rest of the group.

7. Another measure of variation is the interquartile range (IQR) that is the third quartile (Q3) minus the first quartile (Q1). Recall that Q1 is the median of the data points strictly below the median of the distribution. Q3 is the median of the data points strictly above the median of the distribution. Note that the IQR focuses on the middle 50% of a distribution, whereas the range measures the entire distribution from lowest to highest. Have your students calculate Q1, Q3, and the IQR for each group. Referring to the IQRs, discuss how the variations in the groups compare. Also, discuss how conclusions about variation might differ depending on whether the IQR or the range is used. See Table 4.2.2 for a summary of the calculations.

Table 4.2.2 Summary Measures for Each Group

	Min	Max	Range	Q1	Q3	IQR
Two-Digit Group	19.6	25.6	6.0	20.6	23.1	2.5
Three-Digit Group	24.0	29.5	5.5	24.3	26.4	2.1
Four-Digit Group	25.2	34.3	9.1	26.8	31.2	4.4

8. Have students construct side-by-side boxplots. See Figure 4.2.4.

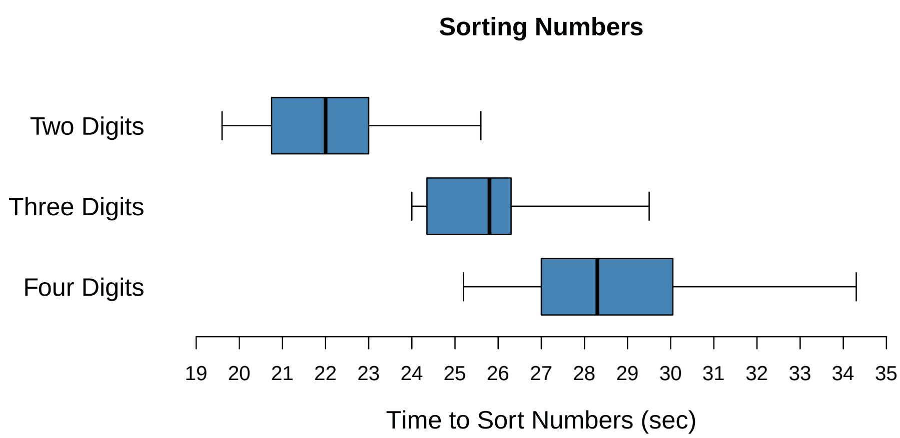


Figure 4.2.4 Side-by-side boxplots of the example class data

9. Ask your students what observations they can make from the boxplots. In particular, is their median measure of center reflected in the boxplots as well as their measures of spread, range, and interquartile range? Discuss how. Note that the boxplots make it clear that the medians are increasing, that the IQR in 2-digit and 3-digit are similar, and that IQR

how far apart two medians are in terms of the number of units of a measure of variability such as IQR. The two distributions being compared have to be of similar variability, and it is the common value that is used to measure how far apart the centers are. Have your students compare the 2-digit and 3-digit distributions. Recall that the median of the 2-digit data set is 22.0 sec and the median of the 3-digit data set is 25.8 sec. The two IQRs are 2.5 and 2.1, which are fairly close. Let's be conservative and take the maximum 2.5 to represent the common spread of the two distributions. By how many IQRs of 2.5 sec do the medians 22.0 and 25.8 differ? The medians of the 2-digit and 3-digit data sets differ by (25.8 – 22.0) / 2.5= 1.5 IQRs.

✏ Interpret the Results in the Context of the Original Question

Have your students recall the original question, "Does the time it takes for a deck of digit cards to be sorted vary with the number of digits in the numbers?" Have your students write a summary of the experiment based on the data collected and analyzed that answers the original question (i.e., what group do they think sorted the cards the fastest and why). They need to support their answer by including the following:

a. A discussion of the plan they used to collect the data
b. The graphs they drew and conclusions made from looking at them
c. The measures of center and variability they computed
d. What the measures said about the comparison of the groups (e.g., whether the measures were similar from group to group).

Example of 'Interpret the Results'

Note: The following is not an example of actual student work, but an example of all the parts that should be included in student work.

We investigated how fast it took us to sort cards that had two-, three-, or four-digit numbers on them. There were 17 cards in each group. We were assigned to one of the groups. To avoid introducing bias into the experimental procedure, we put all our names in a container and then drew them out randomly, one at a time, assigning the first name to the two-digit group, the saw many comparisons such as all the 3-digit and 4-digit times were longer than 75% of the 2-digit times. The median of 3-digit exceeded all 2-digit. So, overall, it was clear that the times to sort the cards are longer as the number of digits in the numbers increases.

It was interesting that the medians (22.0, 25.8, and 28.3) were about the same as the means (22.1, 25.8, 28.7) even though the distributions had all those gaps. We guessed the possible outliers kind of balanced out the distributions. We checked to see if the outliers we saw in the stemplots were also outliers by the 1.5*IQR calculation for boxplots and none were. Different graphs illustrate different things. Finally, we compared the medians of the 2-digit and 3-digit groups by calculating how many common IQRs separated them. We used the maximum IQR of 2.5 for the value of the IQRs and saw that the medians 22.0 and 25.8 differed by (25.8 – 22.0)/2.5 = 1.5 IQRs. We don't really have a number to compare 1.5 to, but it seems to us that 1.5 IQRs is large enough to say the means differ from each other, since they are really separated when we look at the boxplots.

Assessment with Answers

A class of sixth-grade students conducted an experiment involving LEGO blocks to compare the effect of the type of directions provided to a student on the time needed to complete a task. The task was to build a tower from a given set of blocks. A bag of LEGO blocks contained one of the following three sets of directions:

Directions Set 1: Construct a tower using all the blocks in this bag. The longest blocks should be on the bottom and go up in order to the shortest LEGO blocks at the top.

Directions Set 2: Construct a tower using all the blocks in this bag according to the picture. (Figure 4.2.5)

Directions Set 3: Build a tower with the blocks.

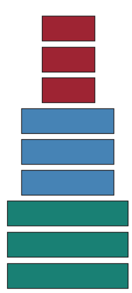
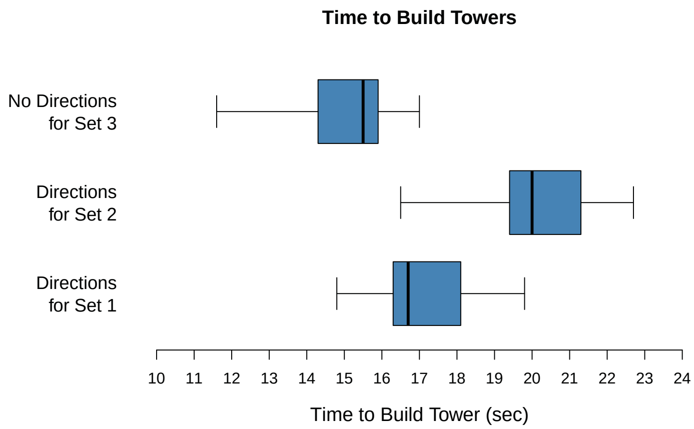


Figure 4.2.5 Diagram shown on directions for set 2

4. Construct side-by-side boxplots of the three groups.



5. Which of the three groups was able to build the tower faster? Using words, numbers, and graphs, explain why you chose the group you did.
Group 3 was able to build the tower the fastest. The median of this group is less than the other two. About 75% of the times for Group 3 are less than all of Group 2 times and 75% of Group 1.

Extension

1. Vary the background of the cards. Using a standard deck of playing cards, create three stacks. Each stack contains the cards ace to 10 with one stack having cards that are all of the same suit, one stack having cards from the two black suits, and one stack having mixed red and black suits. Students would investigate the statistical question, "Does the mixture of suits of cards relate to the amount of time needed to place the cards in order?"

2. Consider Step 11 of the Analysis of the original question in this investigation. Instead of calculating how many common IQRs separate two medians, the separation between means also can be calculated in terms of MADs. Note that the MADs for each group have to close in value so a common value can be determined. Ask your students to calculate mean absolute deviations for the three groups to see if any are similar and, if so, to do Step 11 using MAD in place of IQR.