Section 4

Comparing Groups

Investigation 4.1
How Far Can You Jump?

Overview

This investigation focuses on students conducting a **comparative experiment** to explore the effect a fixed target will have on the distance students can jump from a starting line. Students will be randomly assigned to one of two groups. The first group will be asked to jump as far as they can from the starting line with no target in front of them. The second group will be asked to jump as far as they can, but a target (strip of tape) will be placed on the floor in front of them. Students will collect data about the distance jumped by each member of the two groups. They will display the data in a **back-to-back stemplot** or **boxplot**. Analysis of the data will include graphs and calculations of measures of center and spread.

GAISE Components

This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report.* The four components are formulate a statistical question that can be answered with data, design and implement a plan to collect appropriate data, analyze the collected data by graphical and numerical methods, and interpret the results of the analysis in the context of the original question. This is a GAISE Level B activity.

Learning Goals

Students will be able to do the following after completing this investigation:

- Conduct an experiment to investigate a question
- Collect data and organize the results in a back-to-back stemplot (Level A) or side-by-side boxplots (Level B)
- Use the data to answer the question posed

Common Core State Standards for Mathematical Practice

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.

4. Model with mathematics.
6. Attend to precision.

Common Core State Standards Grade Level Content

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.

6.SP.5 Summarize numerical data sets in relation to their context, such as by the following:

a. Reporting the number of observations

b. Describing the nature of the attribute under investigation, including how it was measured and its units of measurement

c. Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered

d. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered

NCTM Principles and Standards for School Mathematics

Data Analysis and Probability

Grades 6–8 Students should find, use, and interpret measures of center and spread—including mean and interquartile range—and discuss and understand the correspondence between data sets and their graphical representations, especially histograms, stemplots, boxplots, and scatterplots.

Materials

- Masking tape
- Meter sticks
- Recording sheets (included on CD)
- Calculators

Estimated Time

1–2 days

Instructional Plan

Note: You may want to involve the physical education teacher in your school for assistance in this activity. This teacher can give suggestions regarding where to set the target line and how to collect the data.

Formulate a Statistical Question

1. Ask your students if they know what a standing long jump is. Has anyone in class done a standing long jump before? Ask one student to demonstrate a standing long jump for the class. (Several short videos demonstrating the standing long jump are available on YouTube.) Share with your students that Norwegian Arne Tvervaag holds the world record for the standing long jump. He jumped 3.71 meters (12' 2.1") on November 11, 1968.

2. Discuss with your students some reasons why one student might jump farther than another. The following are some possible reasons students may come up with: height of a student, boys might jump farther than girls, what shoes they are wearing, whether there is a prize for the longest jump.

3. After students have generated their own ideas, ask them if they think setting a target line might help a student jump farther. This investigation discusses the statistical question, "Will students jump farther if they are given a fixed target in front of them?"

Collect Appropriate Data

1. Before collecting data, there are procedures that need to be discussed with your students. It is important that your students are placed randomly into a group, that each student performs the jump in the same manner, and that the length of each jump is measured in the same way.

2. The generally accepted way to perform the standing long jump is to 1) stand with both feet up to the start line, 2) take a jump forward with both feet as far as you can, and 3) stay on your feet. **Note:** To avoid injury, this is best done on a mat or grass, instead of a hard floor.

3. The length of the jump should be measured from the start line to the part of the body that lands closest to the start line.

4. Ask students how the two groups should be formed. Students might suggest that there should be an equal number of boys and girls in each group, and some students will want to make sure the best athletes in class are spread between both groups. However, these designs do not ensure randomness. It is important that the groups are formed in a random manner. Random selection helps ensures that the two groups are similar in any attributes that might make a difference in performing the standing long jump. Discuss with your students how you might assign them randomly. One way to select students randomly is to write each of their names on an index card and then, after thoroughly mixing, draw one card at a time from the bag. The student named on the first card is assigned to the No Target group; the student named on the second card drawn is assigned to the Target group. Assignment of students continues to alternate until all the names have been drawn.

5. Set up two stations (one with No Target and one with a Target line) on the playground or in the gym where your students will perform the standing long jump. For the Target group, you may wish to ask the physical education teacher approximately how far your students will be able to jump. You want to set the target line toward the upper limit of what most students can jump. A suggestion for 12-year-olds is 200 cm from the start line.

6. Each student in the No Target group will be asked to jump as far as she/he can from the starting position marked with tape on the floor. Following the jump, with a piece of masking tape, mark the location of the student's heel, or their hand if they fall backward. The heel or hand that is closest to the starting position should be used. Measure the distance in centimeters from the starting point to the end of the jump using a meter stick or extendable tape measure. Record the measurements on the data collection sheet. Similarly, each child in the Target group will be asked to jump as far as she/he can from the starting position marked with tape on the floor. Follow the same procedures as with the No Target group for marking, measuring, and recording the jump.

7. Collect the class data. Display each of the individual student results on the board under the headings No Target group and Target group. An example is shown in Table 4.1.1.

Table 4.1.1 An Example of Data Collected from a Group of 12-Year-Olds

Length in Centimeters for No Target Group														
146	190	109	181	155	167	154	171	157	156	128	157	167	162	137
Length in Centimeters for Target Group														
199	167	147	180	185	170	171	139	154	126	179	158	181	152	

Note that the statistical design being followed is an independent groups one, in which each student participates in exactly one of the two treatments. Is this the best procedure to follow in the context of this problem? Be sure to read the extension and discuss it with your students after the experiment has been completed.

Analyze the Data

1. With the class data displayed on the board, ask your students if they think one group was able to jump farther than the other. Explain to your students that it is difficult to compare groups by just looking at the numbers; it is helpful to organize the data in a graph.

2. Have your students construct a back-to-back stemplot of the results. See Figure 4.1.1. On the board, label the No Target group on the left and the Target group on the right. The stems of the plot are the numbers 10–19, which represent 100 to 190. The "leaf" in the display represents the ones digit.

Jumping Length

No Target		Target
9	10	
	11	
8	12	6
7	13	9
6	14	7
7 6 7 4 5	15	4 8 2
2 7 7	16	7
1	17	0 1 9
1	18	0 5 1
0	19	9

Key: 16|7 represents 167 cm

Figure 4.1.1 Back-to-back stemplot comparing length of jumps for No Target group and Target group

3. Ask your students to modify their back-to-back stemplots showing the data (units digits) ordered. Figure 4.1.2 shows the back-to-back stemplot with the digits in order.

Jumping Length

No Target		Target
9	10	
	11	
8	12	6
7	13	9
6	14	7
7 7 6 5 4	15	2 4 8
7 7 2	16	7
1	17	0 1 9
1	18	0 1 5
0	19	9

Key: 16|7 represents 167 cm

Figure 4.1.2 Back-to-back stemplot comparing length of jumps for No Target group and Target group with the digits in order

4. Ask your students to compare the shapes of the two distributions from the stemplots. Note that the jump lengths in the No Target group are

concentrated between 150–170 cm, whereas those in the Target group are spread out a bit more and appear to be higher in length. The shape of the No Target distribution is peaked, while the shape of the Target distribution is more flat, uniform. There is a gap in the No Target group, suggesting that 109 cm might be what is called an *outlier*, an atypical value. The presence of an outlier might influence the most appropriate measure of center for the data set.

5. Ordering the digits in a stemplot is helpful when finding the quartiles (note the median is the second quartile). The three quartiles are used to construct another graph—the boxplot. To construct a boxplot, have your students find the five-number summary—minimum value, first quartile (Q1) that is the median of the data points strictly below the median of the distribution, the median, the third quartile (Q3) that is the median of the data points strictly above the median of the distribution, and the maximum value. Table 4.1.2 shows the five-number summary for both the Target group and No Target group. Figure 4.1.3 shows the side-by-side boxplots for the data in this example.

Table 4.1.2 Five-Number Summary for Target and No Target Group

	Min	Max	Median	Q1	Q3
No Target Group	109	190	157	146	167
Target Group	126	199	168.5	152	180

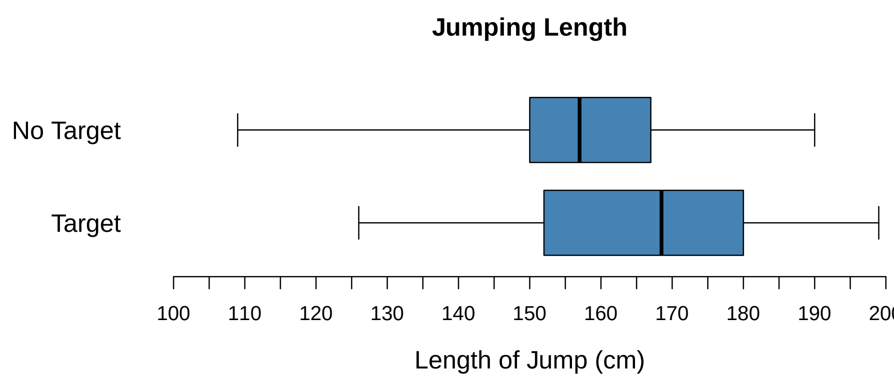


Figure 4.1.3 Side-by-side boxplots comparing length of jumps for No Target group and Target group

6. Remind your students that they are investigating whether a target helps or hinders the length of jumps. Ask your students to discuss several comparisons based on the two boxplots that will contribute to their final answer for the statistical question, “Will students jump farther if they are given a fixed target in front of them?” It is important to have your

students first discuss the meaning of the two boxplots. They should focus on comparing the medians, quartiles, and four sections of the boxplots. Note that in addition to the comparison of shapes they have made, they should note that the median for the Target group is 11.5 cm higher than that for the No Target group. That's a considerable distance. A related note to that comparison of medians is that although the first quartiles are somewhat similar (meaning that 75% of the students in each group jumped at least somewhere around 150 cm), half the students in the Target group jumped more than 168.5 cm, but half the students in the No Target group jumped no more than 157 cm, 11.5 cm shorter. Even more telling is that half the Target group jumped farther than 75% of the No Target group (Target group median is 168.5, No Target group Q3 is 167).

7. In addition to graphing and finding the median and quartiles, ask your students to find another measure of center—the mean length of the jumps. Table 4.1.3 (template available on the CD) shows the sample data and five-number summary and the mean. Discuss with your students whether to use the mean or median. The median is more robust in that it is not influenced by extreme values. The mean is influenced by extreme values, but includes all the information in the calculation. In this example, it appears that 109 is an extreme value in the No Target Group, so the median might be a better measure of center than the mean for the No Target group. Note that whichever measure is used, it should be the same for comparison purposes.

Table 4.1.3 Example Recording Sheet

Student Number	Group 1 - No Targeted Jump (cm)	Group 2 - Targeted Jump (cm)
1	146	199
2	190	167
3	109	147
4	181	180
5	155	185
6	167	170
7	154	171
8	171	139
9	157	154
10	156	126
11	128	179
12	157	158
13	167	181
14	162	152
15	137	
Summary Measures		
Mean	155.8	164.8
Median	157	168.5
Minimum	109	126
Maximum	190	199
Q1	146	152
Q3	167	180

8. Statistics is the study of variability, so a measure of spread needs to be computed to better compare the two groups. Discuss with your students that they calculated one measure of variability when they drew their boxplots, the interquartile range (IQR). The IQR = Q3 – Q1, the difference between the 1st quartile and the 3rd quartile. The IQR provides a measure of the spread of the middle 50% of the jump lengths. In the example data, the IQR of the No Target group is 21 and the IQR for the Target group is 28. This means the middle 50% of the jump lengths for the Target group has a greater spread than the middle 50% of the jump lengths for the No Target group. Discuss with your students what conclusion can be drawn about a data set concerning how spread out it is. Note that a compact data set makes its center more believable that it is reflecting the true value, whereas a widely dispersed data set makes us less sure the center is really characterizing typical performance.

   Have your students compare the two IQRs in words in the context of the data (i.e., what do the IQRs say about how spread out the jump lengths are in the No Target group compared to the Target group). Have them provide a possible contextual explanation as to why they are different. Suggestions will vary. One possibility is that in the presence of a target, people react differently. Some tense up and others push themselves beyond their normal performance.

9. Recall that from the stemplot for the No Target group, 109 was thought to be a possible outlier because it was separated from the rest of the data by a gap. The boxplot allows for a more formal determination as to whether a value should be labeled an outlier (extreme value). The procedure is to calculate what are called the upper fence and lower fence. Data points outside the fences are considered outliers (i.e., data atypical to the data set). The upper fence is Q3 + 1.5*IQR; the lower fence is Q1 – 1.5*IQR. Ask your students to calculate the fences for the No Target group. Note that the lower fence is Q1 – 1.5*IQR = 146 – 1.5*(167 – 146) = 114.5. So, it can be concluded that 109 is an outlier. The implication of this is that, in a statistical analysis of this No Target data set, it would be advisable to use the median as a measure of its center, rather than the mean.

Interpret the Results in the Context of the Original Question

1. Have your students recall the original question, "Will students jump farther if they are given a fixed target in front of them?" Ask your students to write a summary of the experiment that starts with stating an answer

to the question and then supporting their answer with their analysis. They should focus their summary on using center and spread measures, but also include a discussion about the shapes of the graphs they drew.

2. Have your students describe what they think the distribution of jumps with and without a target would be if 2nd graders performed the experiment. Do they think their conclusion they reached about the effect of a target line will be the same for the 2nd graders?

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 conducted a comparative experiment in which some students did a standing long jump with no target in front of them and others did a standing long jump with a target 200 cm in front of them to answer the statistical question, "Will students jump farther if they are given a fixed target in front of them?" (Our gym teacher suggested 200 cm would be a good target for 12-year-olds.)

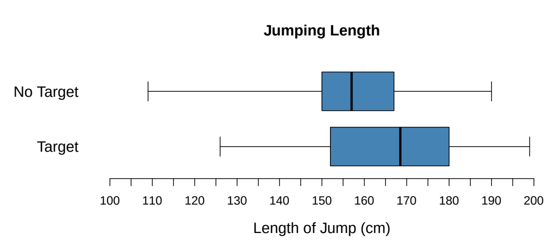
To determine which of us would be in the No Target group and which would be the Target group, we put our names in a hat. The first name randomly drawn from the hat was assigned to the No Target group. The second name drawn was assigned to the Target group. We went back and forth like that until everyone had been assigned to a group.

We measured our distances in centimeters from the starting line to where the closer heel of our shoes landed to the start line. (Everyone landed on their feet.) We tried to make sure everyone did the jump the same way to avoid introducing any sort of bias, like measurement bias, into our results. We drew two comparative graphs of our data.

Jumping Length

No Target		Target
9	10	
	11	
8	12	6
7	13	9
6	14	7
7 7 6 5 4	15	2 4 8
7 7 2	16	7
1	17	0 1 9
1	18	0 1 5
0	19	9

Key: 16|7 represents 167 cm



From the stemplot—except for one possible outlier (109) in the No Target group, it looked like the data sets were spread about the same. But the IQR for the No Target group is 21 and a larger 28 for the Target group, so the middle 50% of the No Target group data is more compact than for the Target group.

Actually, it's better for a data set to have a small variation because it makes us more confident about the centering value. We thought the target group should be more compact because those jumpers had something to concentrate on, but it didn't turn out that way. Regarding the 109, it is an outlier looking at the gap in the stemplot, and it is also an outlier using the Q1 - 1.5*IQR rule for the boxplot. Any value below 146 – 1.5 * (167 – 146) = 114.5 is considered an outlier.

So, did those in the Target group jump farther than the No Target group? From the stemplots, the Target group is shifted to the right compared to the No Target group. Because the No Target group has an outlier, we decided to compare the two groups with medians, rather than means. Based on medians, the answer would be yes, since the median for the Target group was 168.5 cm compared to the median for the No Target group of 157 cm. The Target group jumped a full 11.5 cm longer. In fact, half (seven students) of the Target group jumped farther than 168 cm, but only 3 of the 15 No Target group (20%) jumped that far. Having a target produces higher standing long jump distances. We were wondering if the same conclusion would be made for other age groups. Our guess is that no matter what age groups do this experiment, the results will be similar, since it seems better to have a target as a goal to achieve.

Assessment with Answers

A group of students conducted an experiment to compare the effect of where the target line is placed for the standing long jump. Target lines were placed at 100 cm and 300 cm. Table 4.1.4 shows the length of the jumps in cm for each group.

Table 4.1.4 Jump Lengths (cm) for Groups with Target of 100 cm and 300 cm

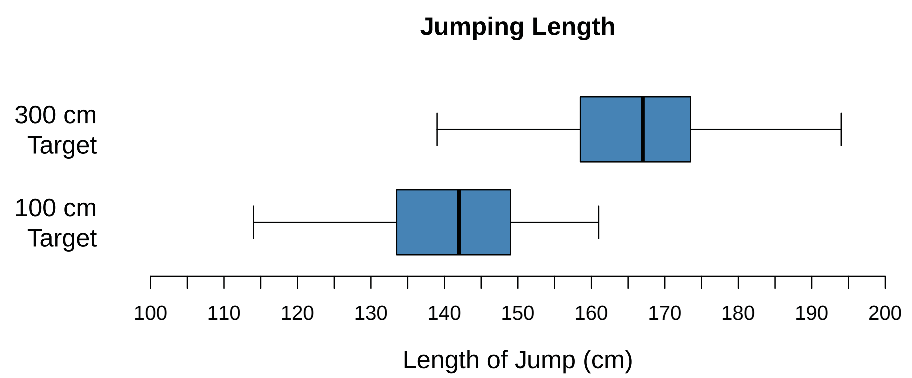
100 cm Target	149	141	161	114	116	142	129	149	138	158	145
300 cm Target	168	185	194	167	147	151	169	178	167	166	139

1. Does the distance a target line is from the start line affect the distance students jump in the standing long jump? Yes, students tended to jump farther when the target line was set at 300 cm.

2. Use words, numbers, and graphs to justify your answer by using at least one graph, a measure of center, and a measure of spread.

Summary

	100 cm target	300 cm target
Mean	140.2	166.5
Minimum	114	139
Q1	129	151
Median	142	167
Q3	149	178
Maximum	161	194
IQR	20	27



Students tended to jump farther when the target line was set at 300 cm than at 100 cm. The mean jumping distance for the 300 cm target was 166.5 cm, while the mean for the 100 cm target was 140.2. The boxplot of the 300 cm target group is shifted much further right than the 100 cm target group. About 75% of the data in the 300 cm target group are greater than about 75% of the 100 cm group.

Extensions

1. As mentioned earlier, the procedure used with all students knowing the experimental condition will no doubt bias the results, as those not assigned to the Target group may imagine a target line. To avoid this potential introduction of bias into the model, redesign the experiment using a **matched pairs** design. Each student does the standing long jump at both stations and the difference—target jump distance minus the no target jump distance—is noted between the two jumps. Your students

should be assigned randomly to which jump they do first. Your students will analyze the differences by making a dotplot, stemplot, or boxplot. If the differences are generally greater than zero, then target jump distances were better than no target distances.

2. Another measure of spread is the mean absolute deviation (MAD), found in Common Core Standard 6.SP.5c (see Investigation 3.4). Calculate the mean absolute deviation (MAD) for each group and compare the two MADs in words in the context of the experiment.

The MAD is the average of the absolute values of the distances from the group's mean. "Deviation" refers to the difference a value is from the mean. "Absolute deviation" is the absolute value of that difference. Column one of Table 4.1.5 contains the data; column two lists the data minus the mean (the deviation); and column three has the absolute value of the deviations in column two. To find the MAD, find the mean of the values in column three.

Table 4.1.5

No Target	No Target – Mean	|No Target – Mean|
146	146 - 155.8 = -9.8	9.8
190	190 - 155.8 = 34.2	34.2
109	109 - 155.8 = -46.8	46.8
181	181 - 155.8 = 25.2	25.2
155	155 - 155.8 = -0.8	0.8
167	167 - 155.8 = 11.2	11.2
154	154 - 155.8 = -1.8	1.8
171	171 - 155.8 = 15.2	15.2
157	157 - 155.8 = 1.2	1.2
156	156 - 155.8 = 0.2	0.2
128	128 - 155.8 = -27.8	27.8
157	157 - 155.8 = 1.2	1.2
167	167 - 155.8 = 11.2	11.2
162	162 - 155.8 = 6.2	6.2
137	137 - 155.8 = -18.8	18.8

The sum of the absolute deviations in this example for the no target data is the sum of the third column, namely 211.6. Dividing the sum by the number of values, 15, yields the mean of 14.1. In words, the average distance away from 155.8 cm that the 15 students jumped was 14.1 cm for the no target group.

Similarly, the MAD for the target group is 16.2 cm. So, according to the point of view of average distance data are from its mean, the target data are spread out more from their mean than the no target data are from their mean. Ask your students if that result is reflected in their boxplots. Why?

References

Franklin, C., G. Kader, D. Mewborn, J. Moreno, R. Peck, M. Perry, and R. Scheaffer. 2007. *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-k–12 curriculum framework*. Alexandria, VA: American Statistical Association. *www.amstat.org/education/gaise*.

National Council of Teachers of Mathematics. 2000. *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Common Core State Standards for Mathematics. www.corestandards.org.

Investigation 4.2
How Fast Can You Sort Cards?

Overview

Students are always interested in how fast they can do something such as playing video games, texting, or running a race. This investigation focuses on the use of a comparative experiment to investigate possible differences in the average time it takes a student to sort a set of 10 cards in numerical order when the size (number of digits) in the numbers varies. Students will be **randomly assigned** to one of three groups. Students in Group 1 will each sort a deck of cards labeled with two-digit numbers. Students in Group 2 will each sort a deck of cards labeled with three-digit numbers. Students in Group 3 will each sort a deck of cards labeled with four-digit numbers. A stopwatch will be used to measure the time needed to complete the task. Students will compare the summary from each group using **measures of center** (mean and median) and **variability** (range, interquartile range, mean absolute deviation) and graphically compare the results using stemplots and boxplots. An informal inference procedure will be introduced as suggested by the Common Core State Standards. This investigation is focused on providing an answer to "Does the time it takes to sort a deck of digit cards vary with the number of digits in the numbers?"

GAISE Components

This investigation follows the four components of statistical problem solving put forth in the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report.* The four components are formulate a statistical question that can be answered with data, design and implement a plan to collect appropriate data, analyze the collected data by graphical and numerical methods, and interpret the results of the analysis in the context of the original question. This is a GAISE Level B activity.

Learning Goals

Students will be able to do the following after completing this investigation:

- Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them
- Explain the idea and use of random assignment
- Conduct an experiment to investigate questions

- Use the data to answer the questions posed
- Collect data and organize the results into stemplots and boxplots
- Compare the results from each group using summary measures of center (such as mean and median) and measures of variability (such as range and interquartile range)

Common Core State Standards for Mathematical Practice

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.

Common Core State Standards Grade Level Content

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.

6.SP.5 Summarize numerical data sets in relation to their context, such as by the following:

a. Reporting the number of observations

b. Describing the nature of the attribute under investigation, including how it was measured and its units of measurement

c. Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.

d. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

7.SP.3 Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities measuring the difference between the centers by expressing it as a multiple of a measure of variability.

NCTM Principles and Standards for School Mathematics

Data Analysis and Probability

Grades 6–8 In grades 6–8, all students should find, use, and interpret measures of center and spread—including mean and interquartile range—and discuss and understand the correspondence between data sets and their graphical representations, especially histograms, stemplots, boxplots, and scatterplots.

Materials

- Three sets of numbered cards (template available on the CD)
- Recording sheets (available on the CD)
- Stopwatches or other timing devices (need to be able to time to nearest 1/10 of a second)

Estimated Time

1–2 days

Instructional Plan

Formulate a Statistical Question

1. Begin the investigation by asking your students when they sort items and what items they sort. Ask if they ever sort numbers in their mathematics class. When finding the median of a set of data, the data must be arranged in order. Tell your students this investigation focuses on sorting cards with numbers on them. Explain that one deck consists of two-digit

numbers, a second with three-digit numbers, and a third with four-digit numbers. Show your students one of these decks of cards. Ask them for factors that may influence how fast they can sort the cards from lowest to highest. They may suggest factors such as the size of the numbers (i.e., the number of digits in the number), the underlying sequence of the numbers, the number of cards, and any incentive offered such as whether there is a prize for the fastest time.

2. Help your students write their suggested factors in the form of a statistical question. This investigation addresses the statistical question, "Does the time it takes to sort a deck of digit cards vary with the number of digits in the numbers?"

✏ Collect Appropriate Data

1. Introduce the idea of comparing the results from three groups of students, each group doing a different version of the task. This is an example of an experiment.

2. Ask students what the variables are in this investigation. Students should realize the first variable of interest is the experimental group (two, three, or four digits) and the second variable of interest is the amount of time needed to complete the task of sorting the cards (as measured in seconds).

3. Discuss with your students the methods they use to select teams on the playground. Are the methods fair? Does each student have the same chance (opportunity) to be selected? What method should we use to assign students to each of the three groups? One way to select students randomly is to write each of their names on an index card and then, after thoroughly mixing, draw one card out of a bag at a time. The student named on the first card is assigned to the Deck 1 Group; the student named on the second card drawn is assigned to the Deck 2 Group; the student named on the third card drawn is assigned to the Deck 3 Group. Assignment of students continues in this pattern until all the names have been drawn.

   Note that some students may suggest that each student roll a die. If a 1 or 2 comes up, the student uses deck 1; 3 or 4, deck 2; 5 or 6, deck 3. Ask them why this method is not desirable. (We should have about the same number of students assigned to the three decks, but it is possible that, in the extreme, all students roll a 5 or 6, say, using this die method.)

4. Point out that in order for them to truly be able to make comparisons, they need to make sure time is measured in the same way for all participants. Therefore, they all need to use the same type of stopwatch and give careful attention to the beginning and ending of the task. Note that one person should do the timing for each student in the specific group (i.e., two-digit, three-digit, or four-digit) to avoid some of the measurement variability.

5. Within each of the three groups, select a member of the group to serve as the timer (leader). Students will perform the sorting task one at a time within each group. Before each student in the group begins, the leader will shuffle the deck of cards (template of cards available on the CD), hand them to a student, and say "GO" and start the stopwatch. The student will sort the cards in ascending order from lowest to highest and say "DONE" when completed. At that time, the leader will stop the stopwatch and record the time on the data collection form.

6. If a student sorts the numbers in the wrong order, the timer should not stop the watch until the numbers are in the correct order from lowest to highest.

7. Collect the class data on the data collection sheet. See Table 4.2.1 for an example.

Table 4.2.1 Example of Class Data

Time (sec) to Sort 2 Digits	Time (sec) to Sort 3 Digits	Time (sec) to Sort 4 Digits
20.6	26.2	31.2
22.9	25.8	28.6
20.9	24.1	28.3
22.2	24.3	31.3
25.6	25.9	26.8
23.1	24.4	27.9
19.6	26.4	28.9
23.6	29.5	27.2
20.5	28.4	34.3
22.0	25.1	26.2
21.8	24.0	25.2

Analyze the Data

1. Begin the analysis by having your students make observations about the class data such as almost all of the data are between 20–30 seconds.

2. Suggest to your students that more observations can be made from graphs. Have your students make a stemplot of each of the three sets of times. See Figures 4.2.1, 4.2.2, and 4.2.3 for examples.

Sort Times for 2-Digit Numbers

19	6
20	5 6 9
21	8
22	0 2 9
23	1 6
24	
25	6

Key: 25|6 represents 25.6 sec.

Figure 4.2.1 Stemplot of sort times for 2-digit numbers

Sort Times for 3-Digit Numbers

24	0 1 3 4
25	1 8 9
26	2 4
27	
28	4
29	5

Key: 26|2 represents 26.2 sec.

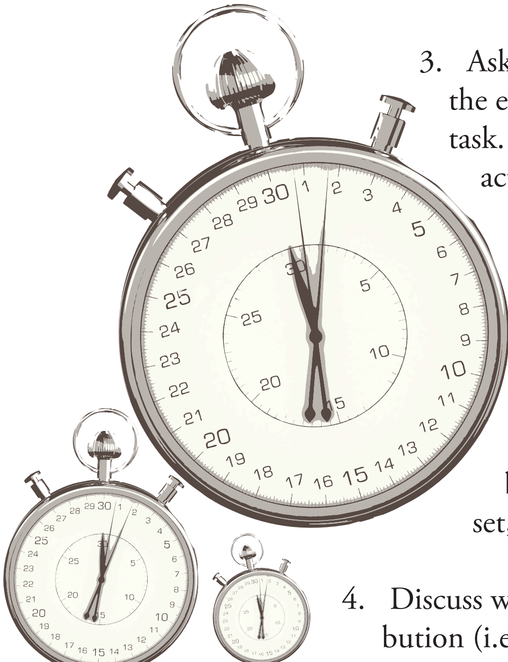
Figure 4.2.2 Stemplot of sort times for 3-digit numbers

Sort Times for 4-Digit Numbers

25	2
26	2 8
27	2 9
28	3 6 9
29	
30	
31	2 3
32	
33	
34	3

Key: 31|2 represents 31.2 sec.

Figure 4.2.3 Stemplot of sort times for 4-digit numbers



3. Ask your students for some observations from the plots regarding the effect the number of digits has on the time to do the ordering task. Note that they should compare shapes: 2-digit might be characterized as bi-modal, 3-digit as "ski-sloped" skewed to the right, and 4-digit as kind of mound-shaped but with big gaps. Each graph shows gaps, but especially 4-digit. Students have to be careful in that the scales of the stemplots are not the same. Drawing dotplots on the same scale would definitely show that 2-digit is to the left of the other two, with 4-digit drifting to the right. The gaps indicate the presence of potential extreme values called outliers. Outliers need to be identified because they can influence conclusions made about the data set, particularly regarding the center.

4. Discuss with your students methods to summarize the center of a distribution (i.e., what could be a representative time needed to complete the sorting task in each group?). Students should suggest that they could use either the mean or the median. Have them do the calculations and then discuss if one measure is more representative of the center of the data in each group than the other and why they think that way. Note that, for these data sets, the respective medians are 22.0, 25.8, and 28.3; the respective (rounded) means are 22.1, 25.8, and 28.7. The medians and means are very close to each other in each group, so either could be used to measure center. Note that the presence of potential outliers in the data sets did not influence the mean as is often the case. Ask your students to look at each data set to see why the medians and means were comparable.

5. Ask your students to comment on what the means or medians are telling them about the typical time taken to complete the task in each group. Note that it's clear the two-digit group is, on average, the quickest, followed by the three-digit group and the four-digit group coming in the slowest.

6. Ask your students whether the overall distributions are the same since their means and medians are about the same in each case. Discuss with them that distributions are compared by their centers and variability. Discuss ways to measure the variability in the data. The range is a basic measure of spread. Recall that the range is the maximum value minus the minimum value. For these data, respectively, the ranges are 6.0, 5.5, and 9.1. Have your students discuss that the first two groups are somewhat similar in how spread out their data are, whereas the third group contains considerably more spread. Looking at the actual data in the stemplots,

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.

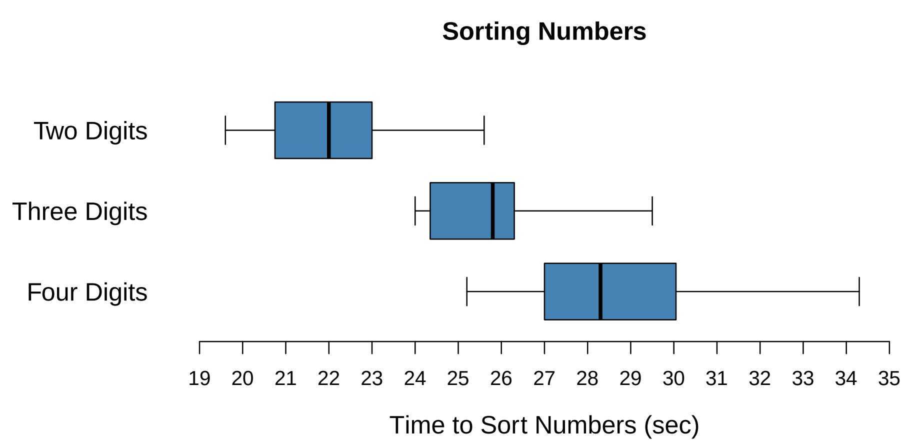


Figure 4.2.4 Side-by-side boxplots of the example class data

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

for 4-digit is about twice as much. All the 3-digit and 4-digit values were higher than 75% of the 2-digit. Seventy-five percent of the 4-digit were higher than 75% of the 3-digit. The 3-digit median exceeded all the 2-digit times.

10. Have your students look at the two types of graphs they have constructed—stemplots and boxplots—and discuss what each of the plots reveal and don't reveal about the comparison of the groups. Lead them to the discovery that several types of graphs should be displayed in a statistical investigation, since each looks at a set of data from a different point of view. Putting all the information together enables the viewer to get a more complete understanding of the experimental results. For example, ask your students if the potential outliers as indicated by the gaps in the stemplots are evident in the boxplots. In boxplots, to identify potential outliers, two calculations need to be made. They are called the lower fence and the upper fence. Values in the data set outside the fences are identified as outliers. The lower fence is Q1 – 1.5*IQR, and the upper fence is Q3 + 1.5*IQR. Have your students calculate the fences for each data set and determine if there are any outliers according to this rule. There are none.

Group	Q1	Q3	IQR	1.5*IQR	Lower Fence	Upper Fence	Outliers
2-digit	20.6	23.1	2.5	3.75	16.85	26.85	None
3-digit	24.3	26.4	2.1	3.15	21.15	29.55	None
4-digit	26.8	31.2	4.4	6.60	20.20	37.80	None

11. In statistical inference, to determine if the centering points of two distributions are statistically close or far apart, their difference is written in terms of the number of units of some measure of variation. Then, that number of units is determined by various techniques to conclude whether the difference of means is small or large (statistically significant). There is a technique your students will be doing as part of the Common Core State Standard in statistics and probability for all high-school students. (There is a formal technique that students who take Advanced Placement Statistics will learn.)

    The Common Core State Standard 7.SP.3 introduces middle-school students to an informal inference procedure by having them measure 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 second to the three-digit group, and the third to the four-digit group. We repeated this until everyone was assigned. After getting our data, we drew stemplots and boxplots.

Sort Times for 2-Digit Numbers

19	6
20	5 6 9
21	8
22	0 2 9
23	1 6
24	
25	6

Key: 25|6 represents 25.6 sec.

Sort Times for 3-Digit Numbers

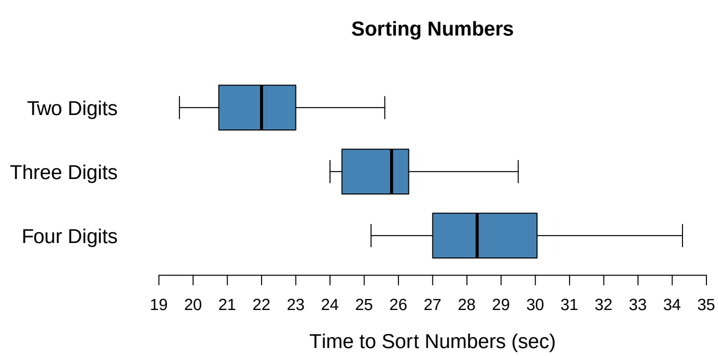
24	0 1 3 4
25	1 8 9
26	2 4
27	
28	4
29	5

Key: 26|2 represents 26.2 sec.

Sort Times for 4-Digit Numbers

25	2
26	2 8
27	2 9
28	3 6 9
29	
30	
31	2 3
32	
33	
34	3

Key: 31|2 represents 31.2 sec.



Each stemplot had at least one gap, indicating there were possible outliers. The two-digit shape had a dip in the middle, but looked symmetric. The three-digit shape was definitely skewed to the right. The four-digit one looked like a triangle for the lower values and then had a couple big gaps. We should have put the stemplots side by side on the same scale like we did with the boxplots. It was really clear from the boxplots that the medians increased and the spread of the middle 50% measured by IQR of the 2-digit and 3-digit data sets was similar, with the spread of the 4-digit about twice as much. We