Georgia Standards of Excellence Curriculum Frameworks

Mathematics

GSE Grade 6

Unit 6: Statistics

Georgia Department of Education

Richard Woods, Georgia's School Superintendent

"Educating Georgia's Future"

Unit 6
Statistics

Table of Contents

OVERVIEW

In this unit students will:
Analyze data from many different sources such as organized lists, box-plots, bar graphs, histograms and dot plots.

- Understand that responses to statistical questions may vary.
- Understand that data can be described by a single number.
- Determine quantitative measures of center (median and/or mean).
- Determine quantitative measures of variability (interquartile range and range).

STANDARDS FOR MATHEMATICAL PRACTICE

1. **Make sense of problems and persevere in solving them.** Students will make sense of the data distributions by interpreting the measures of center and variability in the context of the situations they represent.

2. **Reason abstractly and quantitatively.** Students reason about the appropriate measures of center or variability to represent a data distribution.

3. **Construct viable arguments and critique the reasoning of others**. Students construct arguments regarding which measures of center or variability they would use to represent a particular data distribution. They may critique other students' choices when considering how outliers are handled in each situation.

4. **Model with mathematics.** They use measures of center and variability and data displays (i.e. box plots and histograms) to draw inferences about and make comparisons between data sets. Students need many opportunities to connect and explain the connections between the different representations. Students collect data regarding real-world contexts and create models to display and interpret the data.

5. **Use appropriate tools strategically.** Students consider available tools (including estimation and technology) when answering questions about data or representing data distributions. They decide when certain tools might be helpful. For instance, students in grade 6 may decide to represent similar data sets using dot plots with the same scale to visually compare the center and variability of the data.

6. **Attend to precision.** Students use appropriate terminology when referring data displays and statistical measures.

7. **Look for and make use of structure.** Students examine the structure of data representations by examining intervals, units, and scale in box plots, line plots, histograms and dot plots.

Look for and express regularity in repeated reasoning. Students recognize typical situations in which outliers skew data. They can explain patterns in the way data is interpreted in the various representations they study throughout this unit.

STANDARDS FOR MATHEMATICAL CONTENT

Develop understanding of statistical variability.

MGSE6.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. For example, "How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.

MGSE6.SP.2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.

MGSE6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

Summarize and describe distributions.

MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

BIG IDEAS

- Recognize that statistical questions and the answers account for variability in the data.
- Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
- Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

- Understand that numerical data can be displayed in plots on a number line, including dot plots, histograms, and box plots.
- Summarize numerical data sets in relation to their context, such as by:
  - Reporting the number of observations.
  - Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.
  - Giving quantitative measures of center (median and/or mean) and variability (range or interquartile range), 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.
  - 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

ESSENTIAL QUESTIONS

- What is the best way to organize a set of data?
- What kinds of graphs will best represent a given set of data?
- How can I describe the center of a set of data?
- How can I decide which measure of center (i.e., mean or median) best describes the data?
- How can I describe the spread of a set of data?
- How can I use data to compare different groups?
- How do I choose and create appropriate graphs to represent data?
- What conclusions can be drawn from data?
- How can I recognize when a question is statistical and when it is not?
- What is the difference in a measure of center and a measure of variation?

CONCEPTS & SKILLS TO MAINTAIN

In order for students to be successful, the following skills and concepts need to be maintained:

- Analyzing patterns and seeing relationships
- Fluency with operations on multi-digit numbers and decimals

FLUENCY

It is expected that students will continue to develop and practice strategies to build their capacity to become fluent in mathematics and mathematics computation. The eventual goal is automaticity with math facts. This automaticity is built within each student through strategy development and practice. The following section is presented in order to develop a common understanding of the ideas and terminology regarding fluency and automaticity in mathematics:

Fluency: Procedural fluency is defined as skill in carrying out procedures flexibly, accurately, efficiently, and appropriately. Fluent problem solving does not necessarily mean solving problems within a certain time limit, though there are reasonable limits on how long computation should take. Fluency is based on a deep understanding of quantity and number.

Deep Understanding: Teachers teach more than simply "how to get the answer" and instead support students' ability to access concepts from a number of perspectives. Therefore students are able to see math as more than a set of mnemonics or discrete procedures. Students demonstrate deep conceptual understanding of foundational mathematics concepts by applying them to new situations, as well as writing and speaking about their understanding.

Memorization: The rapid recall of arithmetic facts or mathematical procedures. Memorization is often confused with fluency. Fluency implies a much richer kind of mathematical knowledge and experience.

Number Sense: Students consider the context of a problem, look at the numbers in a problem, make a decision about which strategy would be most efficient in each particular problem. Number sense is not a deep understanding of a single strategy, but rather the ability to think flexibly between a variety of strategies in context.

Fluent students:

- flexibly use a combination of deep understanding, number sense, and memorization.
- are fluent in the necessary baseline functions in mathematics so that they are able to spend their thinking and processing time unpacking problems and making meaning from them.
- are able to articulate their reasoning.
- find solutions through a number of different paths.

For more about fluency, see: http://www.youcubed.org/wp-content/uploads/2015/03/FluencyWithoutFear-2015.pdf and: http://joboaler.com/timed-tests-and-the-development-of-math-anxiety/

STRATEGIES FOR TEACHING AND LEARNING

The purpose of this unit is to begin the study of statistics, beginning with examples of one-variable (numerical) data sets, and evaluating the mean and median of these data sets. These are both ways of describing a data set numerically. Students should become aware of different methods of organizing data, beginning with frequency tables which will then lead into histograms. Guide students to see the similarity between the shape of the histogram and the

shape of the dot plot for the same data set. Teachers should help set the frequency table intervals (histogram intervals) so that a **skyscraper or pancake effect does not occur**. Encourage students to mark the mean and the median on the histograms so that they understand that both of these values describe the center of the data. Students should begin to use statistical language, such a "**distribution**" to describe the spread of the data. Students should also get into the habit of describing the shape of the distribution as "**single peaked, double peaked, roughly symmetric, or skewed**."

Students will also learn about box plots (often called box-and-whisker plots). Box plots are a visual display of the 5-Number Summary: *minimum, lower quartile Q1, median Q2, upper quartile Q3,* and *maximum.* It should be emphasized that the "box" holds the middle 50% of the data, known as the Inter Quartile Range (IQR).

For describing the center of the data set, is it better to use the mean or the median? This may sound like an obscure technical question, but it really can matter. The short answer is "it depends" - to know which you should use, you must know how your data is distributed. The mean is the one to use with symmetrically distributed data; otherwise, use the median. If you follow this rule, you will get a more accurate reflection of an 'average' value. A new task, called Guinness Outliers, has been included in 2015 that will help students better understand this concept. For more information on choosing appropriate measures of center based on context, see http://www.conceptstew.co.uk/PAGES/mean_or_median.html.

MISCONCEPTIONS

- Students may believe all graphical displays are symmetrical. Exposing students to graphs of various shapes will show this to be false.
- Mode is remembered as the "most" and often students think this means the largest value, not "most frequent".
- Students do not remember to put the numbers in order before finding median.
- Students assume that mean is always the best way to describe a set of data, regardless of the context (McGatha, Cobb, & McClain, 1998; cited in Van de Walle, Karp, & Bay-Williams, 2013).
- Students needs to understand that mean is a redistribution of the data, whereas mode and median are not.
- **Students may think that when data is "skewed to the left" that most of the data is on the left. In fact, the tail of the data is on the left and most of the data is on the right. Students confuse clustering and skewing.**

INSTRUCTIONAL RESOURCES AND TOOLS

- Newspaper and magazine graphs for analysis of the spread, shape and variation of data

- Wet Heads In this lesson, students create stem-and-leaf plots and back-to-back stem-and-leaf plots to display data collected from an investigative activity.

- From the National Council of Teachers of Mathematics, Illuminations: Height of Students in our Class. This lesson has students creating box-and-whisker plots with an extension of finding measures of center and creating a stem-and-leaf plot.

- National Library of Virtual Manipulatives. Students can use the appropriate applet from this page of virtual manipulatives to create graphical displays of the data set. This provides an important visual display of the data without the tediousness of the student hand drawing the display.

- Kader, Gary D. "Means and MADs." *Mathematics Teaching in the Middle School* 4.6 (1999): 398-403. Print.

- Franklin, C., Kader, G., Mewborne, D., Moreno, J., Peck, R., Perry, M., Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (gaise) report: A pre-k-12 curriculum framework.* Alexandria, Virginia: American Statistical Association. Print.

- Bright, G., Brewer, W., McClain, K., and Mooney, E. (2001). ***Navigating Through Data Analysis in Grades 6-8*****.** Reston, VA: NCTM.

- ***Statistics and Probability (Grades 6-9)***. Activities that Integrate Math and Science (AIMS Foundation).

SELECTED TERMS AND SYMBOLS

The following terms and symbols are often misunderstood. These concepts are not an inclusive list and should not be taught in isolation. However, due to evidence of frequent difficulty and misunderstanding associated with these concepts, instructors should pay particular attention to them and how their students are able to explain and apply them.

The definitions below are for teacher reference only and are not to be memorized by the students. Students should explore these concepts using models and real life examples. Students should understand the concepts involved and be able to recognize and/or demonstrate them with words, models, pictures, or numbers.

The websites below are interactive and include a math glossary suitable for middle school children. **Note – At the middle school level, different sources use different definitions. Please preview any website for alignment to the definitions given in the frameworks.**

http://www.amathsdictionaryforkids.com/
This web site has activities to help students more fully understand and retain new vocabulary

http://intermath.coe.uga.edu/dictnary/homepg.asp
Definitions and activities for these and other terms can be found on the Intermath website. Intermath is geared towards middle and high school students.

http://www.corestandards.org/Math/Content/mathematics-glossary/glossary

- **Box and Whisker Plot**- A diagram that summarizes data using the median, the upper and lowers quartiles, and the extreme values (minimum and maximum). Box and whisker plots are also known as box plots. It is constructed from the five-number summary of the data: Minimum, Q1 (lower quartile), Q2 (median), Q3 (upper quartile), Maximum.
- **Distribution** – The arrangement of values that show the spread of the data.
- **Dot Plot** – A statistical chart consisting of data points on a number line, typically using circles.
- **Frequency**- the number of times an item, number, or event occurs in a set of data
- **Grouped Frequency Table**- The organization of raw data in table form with classes and frequencies
- **Histogram**- a way of displaying numeric data using horizontal or vertical bars so that the height or length of the bars indicates frequency

- **Inter-Quartile Range (IQR**)- The difference between the first and third quartiles. (Note that the first quartile and third quartiles are sometimes called upper and lower quartiles.)

- **Maximum value**- The largest value in a set of data.

- **Mean**- The "average" or "fair share" value for the data. The mean is also the balance point of the corresponding data distribution.

$$arithmetic\ mean = \bar{x} = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n}$$

- **Measures of Center**- The mean and the median are both ways to measure the center for a set of data.

- **Measures of Spread**- The range and the interquartile range are both common ways to measure the spread for a set of data.

- **Median**- The value for which half the numbers are larger and half are smaller. If there are two middle numbers, the median is the arithmetic mean of the two middle numbers. Note: The median is a good choice to represent the center of a distribution when the distribution is skewed or outliers are present.

- **Minimum value**- The smallest value in a set of data.

- **Mode**- The number that occurs the most often in a list. There can be more than one mode, or no mode.

- **Numerical Data**- Consists of numbers only. Numerical data can be any rational numbers.

- **Outlier**- An outlier is an observation that is numerically distant from the rest of the data.

- **Range**- A measure of spread for a set of data. To find the range, subtract the smallest value from the largest value in a set of data.

- **Skewed Data** –When a set of data is not symmetrical it can skewed, meaning it tends to have a long tail on the left or right side.

- **Statistical Questions** - A statistical question is one for which you don't expect to get a single answer. Instead, you expect to get a variety of different answers, and you are interested in the distribution and tendency of those answers. For example, "How tall are you?" is not a statistical question, however "How tall are the students in your school?" *is* a statistical question.

- **Variability** – Describes how spread out or closely clustered a set of data is. Variability includes range and interquartile range.

FORMATIVE ASSESSMENT LESSONS (FAL)

Formative Assessment Lessons are intended to support teachers in formative assessment. They reveal and develop students' understanding of key mathematical ideas and applications. These lessons enable teachers and students to monitor in more detail their progress towards the targets of the standards. They assess students' understanding of important concepts and problem solving performance, and help teachers and their students to work effectively together to move each student's mathematical reasoning forward.

More information on Formative Assessment Lessons may be found in the Comprehensive Course Guide.

SPOTLIGHT TASKS

For middle and high schools, each Georgia Standards of Excellence mathematics unit includes at least one Spotlight Task. The Spotlight Tasks serve as exemplars for the use of the Standards for Mathematical Practice, appropriate unit-level Georgia Standards of Excellence, and research-based pedagogical strategies for instruction and engagement. Each task includes teacher commentary and support for classroom implementation. Some of the Spotlight Tasks are revisions of existing Georgia tasks and some are newly created. Additionally, some of the Spotlight Tasks are 3-Act Tasks based on 3-Act Problems from Dan Meyer and Problem-Based Learning from Robert Kaplinsky.

3-ACT TASKS

A Three-Act Task is a whole group mathematics task consisting of 3 distinct parts: an engaging and perplexing Act One, an information and solution seeking Act Two, and a solution discussion and solution revealing Act Three.

More information along with guidelines for 3-Act Tasks may be found in the Comprehensive Course Guide.

TASKS

Task	Task Type *Grouping Strategy*	Content Addressed	Standards Addressed
What is a Statistical Question?	Learning/Scaffolding *Individual/Partner*	Statistical Questions	MGSE.6.SP.1
Who was the Greatest Yankee Homerun Hitter?	Learning / Scaffolding Task *Individual/Partner*	Mean, Median, Mode, Stem and Leaf Charts, Frequency lists, Histograms, Range	MGSE.6.SP.2 MGSE.6.SP.3 MGSE.6.SP.4 MGSE.6.SP.5 MGSE.6.SP.5.a MGSE.6.SP.5.c MGSE.6.SP.5.d
How Long is a Minute?	Learning / Scaffolding Task *Individual/Partner*	Box Plots, Inter Quartile Range (IQR), Minimum, Maximum, 5 number summary	MGSE6.SP.2 MGSE6.SP.3 MGSE6.SP.4 MGSE6.SP.5 MGSE6.SP.5.c MGSE6.SP.5.d
Cost of Learning (Spotlight Task)	Learning Task *Individual/Partner*	Comparing Box Plots	MGSE6.SP.4 MGSE6.SP.5 MGSE6.SP.5.d
Where's Waldo?	Learning / Scaffolding Task *Individual/Partner* Differentiation Option	Side by Side Box Plots; A deeper look into percentages that are present in box plots, discussion of outliers. Differentiation option at end of task	MGSE.6.SP.1 MGSE.6.SP.2 MGSE.6.SP.3 MGSE.6.SP.4 MGSE.6.SP.5 MGSE.6.SP.5.a MGSE.6.SP.5.b MGSE.6.SP.5.c MGSE.6.SP.5.d
Guinness Outliers	Learning / Scaffolding Task	Measures of Center mean vs. median	MGSE.6.SP.4 MGSE.6.SP.5
Mean, Median, Mode and Range (FAL)	Formative Assessment Lesson *Partner/SmallGroup*	Mean, Median and Range	MGSE.6.SP.1 MGSE.6.SP.2 MGSE.6.SP.3
Suzi's Company	Short Cycle Task	Mean, Median, Range and Measure of Center	MGSE.6.SP.4 MGSE.6.SP.5 MGSE.6.SP.5.a MGSE.6.SP.5.b MGSE.6.SP.5.c MGSE.6.SP.5.d

Candy Bars	Short Cycle Task	Analyze Data and Justify Conclusions	MGSE.6.SP.4 MGSE.6.SP.5 MGSE.6.SP.5.a MGSE.6.SP.5.b MGSE.6.SP.5.c MGSE.6.SP.5.d
Culminating Task: Order Up! Fast Food Frenzy Part I: Burgers~ I'm Lovin' It! Part II: Biggie Size those Fries~ or Maybe Not! Part III: I'm Thirsty!	Performance Task *Individual Task*	Part I: Measures of Center, measures of spread, overall shape, outliers Part II: Side by Side Box Plots, Spread, Shape, Variability Part III: Histograms, measures of spread, graphing calculator extension	MGSE.6.SP.2 MGSE.6.SP.3 MGSE.6.SP.4 MGSE.6.SP.5 MGSE.6.SP.5.a MGSE.6.SP.5.b MGSE.6.SP.5.c MGSE.6.SP.5.d

What is a Statistical Question?

Back to Task Table

In this task, students will be presented with a scenario then asked to create a list of questions that could go along with the situation. These questions should be recorded on the board and on a student recording sheet. Students will then determine which questions can be answered with just one answer (i.e. non-statistical questions) versus those that contain variability in their answers (i.e. statistical questions).

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.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. For example, "How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.

STANDARDS OF MATHEMATICAL PRACTICE

2. Reason abstractly and quantitatively. Students reason to determine if a question is statistical or not.
3. Construct viable arguments and critique the reasoning of others. Students construct arguments to prove that a question is statistical. Students will take a question and make it into a statistical question that can be used to collect data.
6. Attend to precision. Students use appropriate terminology when referring to statistical questions and expected.
7. Look for and make use of structure. Students examine the structure of statistical questions and recognize a statistical question as one that anticipates variability in the data related to the question. All statistical questions anticipate variability in the data, have specific populations, and account for the variability in the answers (data collected).
8. Look for and express regularity in repeated reasoning. Students identify possible statistical questions from the graphical display.

ESSENTIAL QUESTIONS

- What is a statistical question?
- How do I write a statistical question?
- How do I recognize if a question is statistical or not?
- How do I write a statistical question by looking at the graphical display of the data?

MATERIALS NEEDED

An interesting scenario to present to the class: Ideas:
• Project a picture of a crowd of people from a recent event such as a festival or sporting event.
• Show a portion of a recent news clip.
• Write a statement such as, "This weekend I went to a basketball game" on the board.

What is a Statistical Question?

1. Examine the situation presented by your teacher.

    a. Make a list of at least 10 questions you could ask about this scenario.

    b. Circle questions that can be answered with a single answer such as, "What is your name?

    c. Underline the questions with answers containing variability such as, "How far do you live from your classmates?"

2. What must be true about a question for it to classified as a statistical question?
All statistical questions anticipate variability in the data, have specific populations, and account for the variability in the answers (data collected).

Students should engage in discussion and sharing of parts 1and 2 before continuing

3. For each question, decide if it is a statistical question; if it is, put an S beside it. If it is not, EXPLAIN why it is not and REWRITE it as a statistical question.

A. How many words are there in this sentence? ***Anticipates a fixed answer, ex. How many words are in each sentence of the book, The Hunger Games.***

B. How many TV's are in your house? ***Anticipates fixed answers, ex. How many TV's are in the homes of the people in your neighborhood?***

C. How many siblings do the students on Mrs. Johnson's homeroom have living in their homes?
S

D. How many socks are in your drawer? ***There will be only one answer. How many socks do the students in the sixth grade class own?***

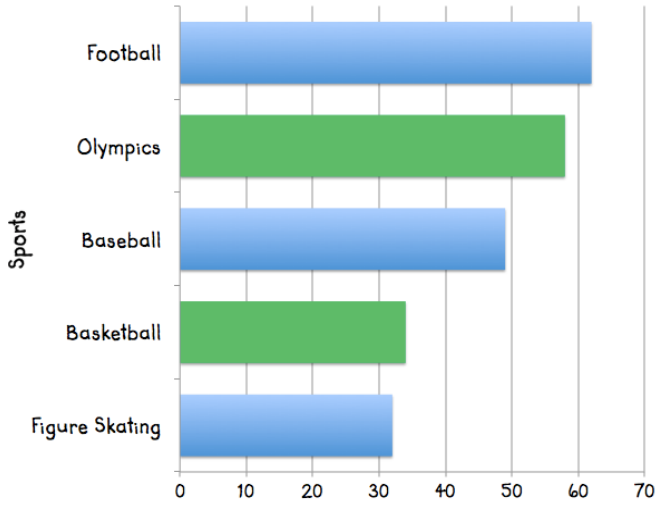
F. What is your favorite color? ***There will be only one answer. What are the favorite colors of all the students that attend Cross Schools?***

G. How far does Savannah have to walk to reach Colin's house each day? ***Anticipates a fixed answer, ex. How far do all of Colin's friends have to walk to reach his house?***

4. Look at each graphical display and write a question that COULD have been asked to collect the specific data.

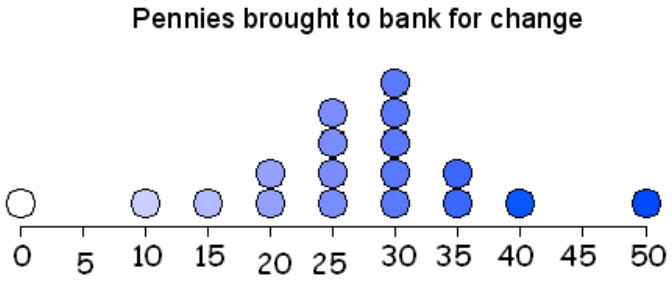
A.

Answers will vary - What are the favorite sports of the students in the sixth grade at Inman Middle School?



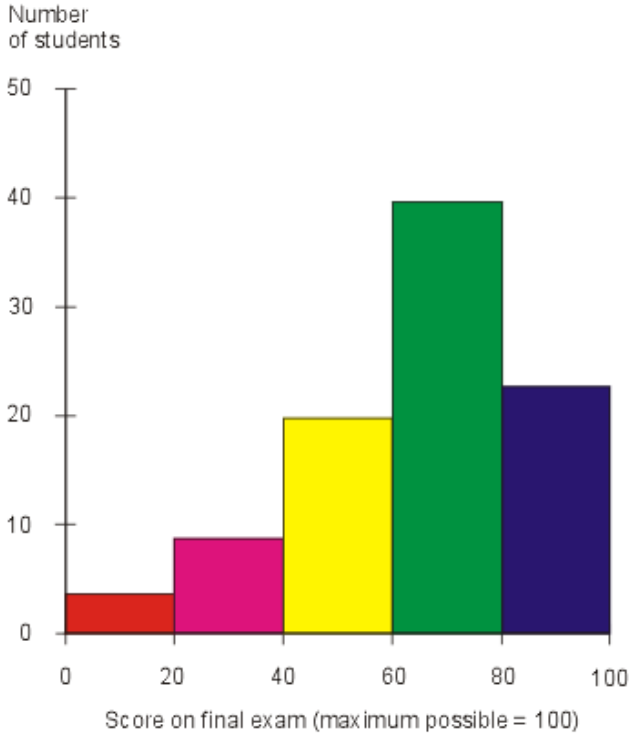
B.

Answers will vary - Mr. Oliver asked all the students to bring in pennies for an activity they were planning. How many pennies did each student bring in for the activity?



C.

Answers will vary - What were the scores on Unit 5 Test of all the students in the sixth grade?



D. (Mark represents any category that the student can make up)

Mark	Tally	Frequency
4	II	2
5	II	2
6	IIII	4
7	𝍸	5
8	IIII	4
9	II	2
10	I	1

Answers will vary - How many pieces of candy did each student on the Harvey Team win during the game?

What is a Statistical Question?

1. Examine the situation presented by your teacher.

    a. Make a list of at least 10 questions you could ask about this scenario.

    b. Circle questions that can be answered with a single answer such as, "What is your name?

    c. Underline the questions with answers containing variability such as, "How far do you live from your classmates?"

2. What must be true about a question for it to classified as a statistical question?

3. For each question, decide if it is a statistical question; if it is, put an S beside it. If it is not, EXPLAIN why it is not and REWRITE it as a statistical question.

A. How many words are there in this sentence?

B. How many TV's are in your house?

C. How many siblings do the students on Team B have living in their homes?

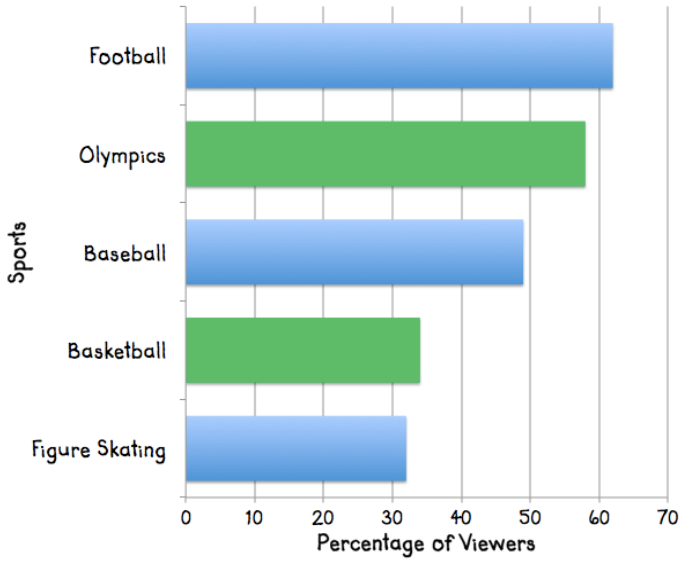
D. How many socks are in your drawer?

F. What is your favorite color?

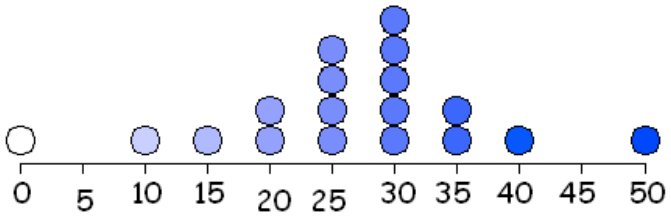
G. How far does Savannah have to walk to reach Colin's house each day?

4. Look at each graphical display and write a question that COULD have been asked to collect the specific data.

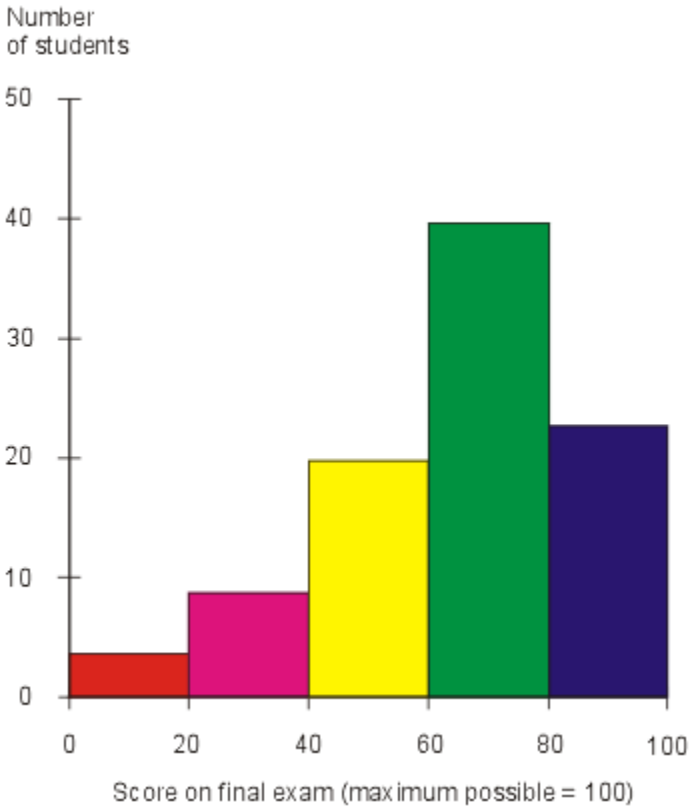
A



B.


0
5
10
15
20
25
30
35
40
45
50

C.


Number
of students
50
40
30
20
10
0
0
20
40
60
80
100
Score on final exam (maximum possible = 100)

D. (Mark represents any category that the student can make up)

Mark	Tally	Frequency
4	||	2
5	||	2
6	||||	4
7	𝍸	5
8	||||	4
9	||	2
10	|	1

Who was the Greatest Yankee Homerun Hitter?

Back to Task Table

In this task, students will analyze historical baseball statistics in order to make inferences and answer questions about the data.

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
MGSE6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.
MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS OF MATHEMATICAL PRACTICES ADDRESSED

1. **Make sense of problems and persevere in solving them.** Students will make sense of the data distributions by interpreting the measures of center and variability in the context of the situations they represent.
2. **Reason abstractly and quantitatively.** Students reason about the appropriate measures of center or variability to represent a data distribution.
3. **Construct viable arguments and critique the reasoning of others.** Students construct arguments regarding which measures of center or variability they would use to represent a particular data distribution. They may critique other students' choices when considering how outliers are handled in each situation.
4. **Model with mathematics.** They use measures of center and variability and data displays (i.e. box plots and histograms) to draw inferences about and make comparisons between data sets. Students need many opportunities to connect and explain the connections between the different representations. Students collect data regarding real-world contexts and create models to display and interpret the data.
6. Attend to precision. Students use appropriate terminology when referring data displays and statistical measures.

ESSENTIAL QUESTIONS

- What is meant by the center of a data set, how is it found and how is it useful when analyzing data?
- What conclusions can be drawn from data?

TEACHER NOTES

Before starting task, show students the tables and ask, "What questions can we write from this information?" Record a list of answers for the class to see (whiteboard, padlet.com, etc.). It is likely that students will ask, "Who is the best player?" to which the teacher can respond with, "How will we know?" Let students brainstorm how they can determine which player is the best hitter before presenting them with the task questions.

TASK: WHO WAS THE GREATEST YANKEE HOME RUN HITTER?

The following table lists four of the greatest New York Yankees' home run hitters with the number of homeruns each hit while a Yankee.

Adapted from: James M. Landwehr and Ann E. Watkins, Dale Seymour Publications, Mathematics, 1986, Pg. 160

Babe Ruth		**Lou Gehrig**		**Mickey Mantle**		**Roger Maris**	
Year	Home runs	Year	Home runs	Year	Home runs	Year	Home runs
1920	54	1923	1	1951	13	1960	39
1921	59	1924	0	1952	23	1961	61
1922	35	1925	20	1953	21	1962	33
1923	41	1926	16	1954	27	1963	23
1924	46	1927	47	1955	37	1964	26
1925	25	1928	27	1956	52	1965	8
1926	47	1929	35	1957	34	1966	13
1927	60	1930	41	1958	42		
1928	54	1931	46	1959	31		
1929	46	1932	34	1960	40		
1930	49	1933	32	1961	54		
1931	46	1934	49	1962	30		
1932	41	1935	30	1963	15		
1933	34	1936	49	1964	35		
1934	22	1937	37	1965	19		
		1938	29	1966	23		
		1939	0	1967	22		
				1968	18		

Source: Macmillan Baseball Encyclopedia, 4th edition

1. Find the mean, median, and number of observations for each player.

	Ruth		*Gherig*		*Mantle*		*Maris*
Mean	*43.9*		*29*		*29.7*		*29*
Median	*46*		*32*		*28.5*		*26*
n	*15*		*17*		*18*		*7*

Of the two values you computed for each player, which do you think best describes the performance of each player? Why?
The median is not affected by extreme high or low values, so the median would best describe the performance of each player.

2. Make a frequency table and histogram for each player. Use the intervals 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69

Comment
GIVING the students the intervals ensures that they will be able to compare. Students should understand that a number cannot be used twice in two different intervals.

It is important to emphasize several points about these plots.

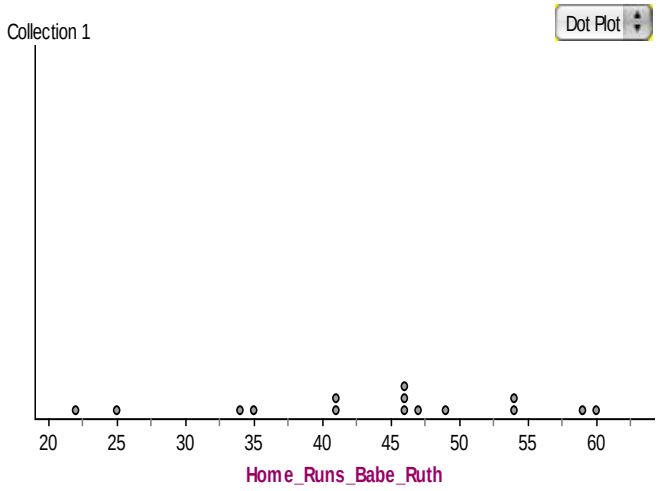
- ***Two of the players, Gehrig and Maris, had years in which their home run totals were single digits. Be sure to point out how single digits are handled in a plot where everything else is a two-digit number.***

3. Describe the shape of the data for all four players. What observations can you make about the four players by looking at the shape?

4-6) Options: Rather than telling students to look more closely at Babe Ruth, allow each group to pick a different player to study. This would allow for summarizing dialogue comparing players and determine who the better hitter is.

4. Looking at the histogram for BABE RUTH, determine the range of homeruns Babe Ruth hit while playing for the Yankees.
60-22=38 homerun range

5. Create a dot plot using Babe Ruth's home runs.



6. Describe the similarities between Babe Ruth's histogram and Babe Ruth's dot plot. Are there any differences? Does the range change between the histogram and the dot plot?
The majority of the data is centered between 40-50. The range of data for both graphs is between 22 to 60.

Who was the Greatest Yankee Homerun Hitter?

The following table lists four of the greatest New York Yankees' home run hitters with the number of homeruns each hit while a Yankee.

Adapted from: James M. Landwehr and Ann E. Watkins, Dale Seymour Publications, Mathematics, 1986, Pg. 160

Babe Ruth		Lou Gehrig		Mickey Mantle		Roger Maris	
Year	Home runs	Year	Home runs	Year	Home runs	Year	Home runs
1920	54	1923	1	1951	13	1960	39
1921	59	1924	0	1952	23	1961	61
1922	35	1925	20	1953	21	1962	33
1923	41	1926	16	1954	27	1963	23
1924	46	1927	47	1955	37	1964	26
1925	25	1928	27	1956	52	1965	8
1926	47	1929	35	1957	34	1966	13
1927	60	1930	41	1958	42		
1928	54	1931	46	1959	31		
1929	46	1932	34	1960	40		
1930	49	1933	32	1961	54		
1931	46	1934	49	1962	30		
1932	41	1935	30	1963	15		
1933	34	1936	49	1964	35		
1934	22	1937	37	1965	19		
		1938	29	1966	23		
		1939	0	1967	22		
				1968	18		

Source: Macmillan Baseball Encyclopedia, 4th edition

1. Find the mean, median, and number of observations for each player.

	Ruth		***Gehrig***		***Mantel***		***Maris***
Mean							
Median							
n							

Of the two values you computed for each player, which do you think best describes the performance of each player? Why?

2. Make a frequency table and histogram for each player. Use the intervals 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69

3. Describe the shape of the data for all four players. What observations can you make about the four players by looking at the shape?

4. Looking at the histogram for BABE RUTH, determine the range of homeruns Babe Ruth hit while playing for the Yankees.

5. Create a dot plot using Babe Ruth's home runs.

6. Describe the similarities between Babe Ruth's histogram and Babe Ruth's dot plot. Are there any differences? Does the range change between the histogram and the dot plot?

How Long is a Minute? (Learning About Box Plots Kinesthetically)

Back to Task Table

In this task, students will engage in a kinesthetic activity to explore how box plots can be used to organize and represent data.

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS FOR MATHEMATICAL PRACTICE

1. Make sense of problems and persevere in solving them. Students will make sense of the data distributions by interpreting the measures of center and variability in the context of the situations they represent.
4. Model with mathematics. They use measures of center and variability and data displays (i.e. box plots and histograms) to draw inferences about and make comparisons between data sets. Students need many opportunities to connect and explain the connections between the different
6. Attend to precision. Students use appropriate terminology when referring data displays and statistical measures.
8. Look for and make use of structure. Students examine the structure of data representations by examining intervals, units, and scale in box plots, line plots, histograms and dot plots.

ESSENTIAL QUESTIONS

- What is meant by the center of a data set, how is it found and how is it useful when analyzing data?
- How can I describe the spread of a set of data?
- What conclusions can be drawn from data?
- What information is necessary to create a box and whisker plot?

MATERIALS

Stop Watch
White Paper
String or yarn
Graphing calculator (optional)

How Long is a Minute? (Learning About Box Plots Kinesthetically)

Do you think you can determine how long a minute without looking at a clock? With your partner and your stop watch, you will each attempt to determine when you think a minute has passed without looking at a watch or clock. All time will be recorded in SECONDS – do not convert to minutes. You will take turns measuring and timing.
Teacher - Partners should sit as far as possible from other students in the class. Make sure students cannot see a clock. Have students remove watches. Each group of 2 students will need a stop watch.

1. Within your pair decide who will be the "timer" and who will be the "guesser." The "timer" will have the stop watch and direct the "guesser" when to start. The "guesser" will attempt to tell how long a minute is without looking at a clock.

2. When instructed by teacher, the "timer" tells the "guesser" to begin. When the "guesser" believes a minute has passed, he should say, "stop" quietly. Timer - record the time that has passed to the nearest second. Do not tell your partner how much time actually passed! The "timer" needs to record the "guesser" time down on a sheet of paper.

3. Switch roles and repeat #2.

4. Share times with your partner. Write **your own** time down (in seconds) on a sheet of paper, large enough so that everyone can see it (one sheet per student).
Teacher – Once everyone is finished have students stand in line from smallest to largest, and hold the pages in front of you, so that everyone can see them.
***Comment: Have students line up from least to greatest. At this point, you may choose to have your students tape their papers to the board, and then sit down in their desks for the remainder of the* lesson.**
Comment: The string can also be taped onto the board around Q1 and Q3.
Optional Tools: Large sticky notes or white board could replace the large piece of paper.

As a class, find the median of the data by counting to the middle. Have a student put a sticky note with "Q2" where the median is. Note to class - The median is also known as Quartile 2 (Q2)
Comment: If the median is in between two values, have both the students that are holding the two values it falls between hold the sticky note for Q2.
Teachers emphasis that the Q2 splits the data in half.

As a class, find the median of the lower half of numbers. Have a student put a sticky note with "Q1" on this place. Note to class - This is Quartile 1 (Q1) the median of the lower half of data.
Comment: If Q1 is in between two values, have both the students that are holding the two values it falls between hold the sticky note for Q1.

As a class, find the median of the upper half of numbers. Have a student put a sticky note with "Q3" on this place. Note to class - This is Quartile 3 (Q3) the median of the upper half of data.
Comment: If Q3 is in between two values, have both the students that are holding the two values it falls between hold the sticky note for Q3.

Put a string around the people from Q1-Q3. Note to class - This is the "box" of the box plot.

ASK STUDENTS: What percent of students lie in this box? 50% of the data set lies within this area.
Then, put strings from the ends of the box to the minimum and maximum. These are the "whiskers" of the box and whisker plot.
ASK STUDENTS: What percent of students lie to the RIGHT of the box? 25% of the data lies in each whisker.
ASK STUDENTS: What percent of students lie to the LEFT of the box? 25% of the data lies in each whisker.

Discuss the Box Plot they have created. NOTE to students that a box plot requires a number line.
Comment: Be sure to record each time down on this paper so you do not forget it! Students will re-write these times down on a blank sheet of paper LARGE enough for the class to see in the next step.

4. From the data draw a box plot of the class's data that is the same box plot that you made with your bodies. Remember to draw the number line FIRST.

5. How many students participated in the live Box Plot? _____

6. How many students were INSIDE the box? _____ What percent is this? _____
Comment: this should be half of the students – 50%

7. Where are the other students?
Comment: 25% are below the box and 25% are above the box

8. What conclusions can you make about the each quartile?
Comment: Each quartile represents 25% of the students

OPTIONAL - Input the data in List 1 in the Graphing Calculator. This is the same box plot you have drawn and created with your bodies. ***Comment: This is done in the Graphing Calculator (TI-84) by inputting the data in L1 under the Stat Menu. Then go to 2nd-StatPlot-Plot 1. Choose the Box Plot in the bottom row in the middle. Press Zoom-Statistics to show the box plot graph.***

How Long is a Minute? (Learning About Box Plots Kinesthetically)

Do you think you can determine how long a minute is without looking at a clock? With your partner and your stop watch, you will each attempt to determine when you think a minute has passed without looking at a watch or clock. All time will be recorded in SECONDS – do not convert to minutes. You will take turns measuring and timing.

1. Within your pair decide who will be the "timer" and who will be the "guesser." The "timer" will have the stop watch and direct the "guesser" when to start. The "guesser" will attempt to tell how long a minute is without looking at a clock.

2. When instructed by teacher, the "timer" tells the "guesser" to begin. When the "guesser" believes a minute has passed, he should say, "stop" quietly. Timer - record the time that has passed to the nearest second. Do not tell your partner how much time actually passed! The "timer" needs to record the "guesser" time down on a sheet of paper.

3. Switch roles and repeat #2.

4. Share times with your partner. Write **your own** time down (in seconds) on a sheet of paper, large enough so that everyone can see it (one sheet per student).

As a class, find the median of the data by counting to the middle. Have a student put a sticky note with "Q2" where the median is. Note to class - The median is also known as Quartile 2 (Q2)

As a class, find the median of the lower half of numbers. Have a student put a sticky note with "Q1" on this place. Note to class - This is Quartile 1 (Q1), which is the median of the lower half of the data.

As a class, find the median of the upper half of numbers. Have a student put a sticky note with "Q3" on this place. Note to class - This is Quartile 3 (Q3), which is the median of the upper half of the data.

Create a human box plot or box plot using class data however the class decides. (Sticky notes and white board or using string for human box plot.)

4. From the data draw a box plot of the class's data that is the same box plot that you made with your bodies. Remember to draw the number line FIRST.

5. How many students participated in the live Box Plot? _____

6. How many students were INSIDE the box? _____ What percent is this? _____

7. Where are the other students?

8. What conclusions can you make about the each quartile?

Cost of Learning (Spotlight Task)

Back to Task Table

Task adapted from Illuminations Box Plotter Activity, http://illuminations.nctm.org/Activity.aspx?id=3541

In this task, students will explore the meaning of box plots in the context of school finance.

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

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.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

ESSENTIAL QUESTIONS

- How are box plots compared?
- How can information be used to better understand our world?

MATERIALS REQUIRED

- Student Handout

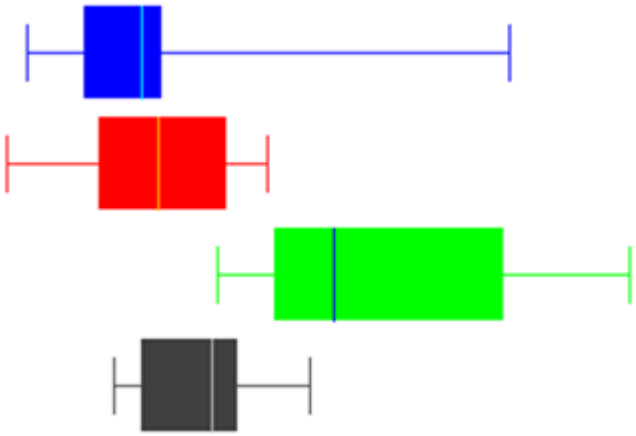
TEACHER NOTES

In this task, students will be presented with four box plots and asked to tell what they notice. They will then be asked to discuss what they wonder or are curious about. These questions will be recorded on a class chart or on the board. Students will then use mathematics to answer their own questions. Students will be given information to solve the problem based on need. When they realize they don't have the information they need, and ask for it, it will be given to them

LAUNCH:

Display or handout the box plots to be compared.

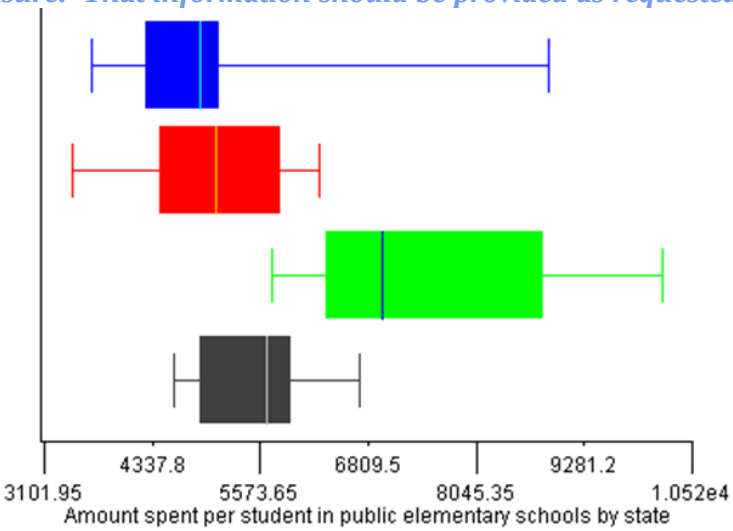
What questions do you have about this information?
What questions can you answer with this information?



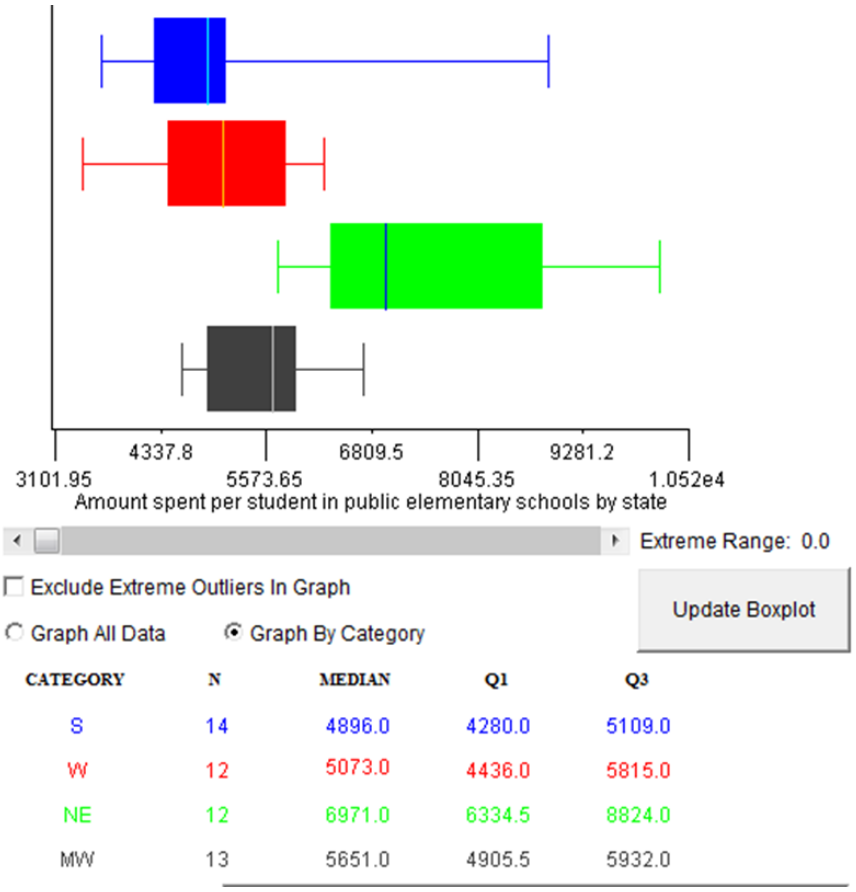
EXPLORE:

Write statistical questions that can be answered with this information. Answer your questions mathematically.

Students should realize the need for more information, including what is being compared and the unit of measure. That information should be provided as requested.



Each box plot represents a different region in the United States. The first box plot represents the south (blue). The second represents the west (red). The third represents the northeast (green) and the fourth represents the midwest (black).



CATEGORY	N	MEDIAN	Q1	Q3
S	14	4896.0	4280.0	5109.0
W	12	5073.0	4436.0	5815.0
NE	12	6971.0	6334.5	8824.0
MW	13	5651.0	4905.5	5932.0

CONNECT/REFLECT:
Students will compare and share their questions and answers.

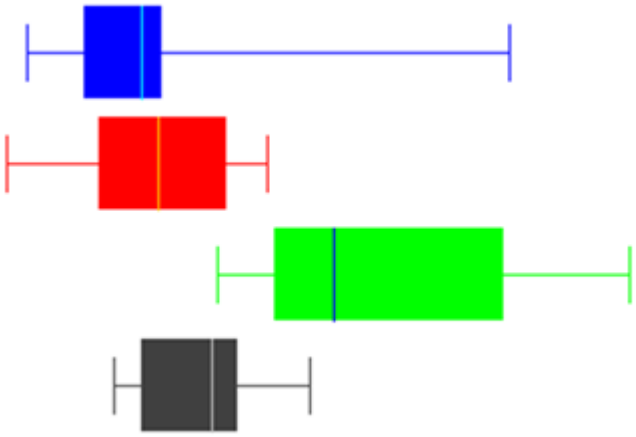
Intervention: Guiding questions for struggling students:

- Which region spent the most/least per student?
- Within each region, which one varies the most/least? How do you know?

Why do you think the amount spent per student varies from region to region?

Cost of Learning

LAUNCH:



What questions do you have about this information?

What questions can you answer with this information?

EXPLORE:

Write statistical questions that can be answered with this information. Answer your questions mathematically.

CONNECT/REFLECT:

How can this information be used to better understand the world around us?

Where's Waldo?

Back to Task Table

In this task, students will use "Where's Waldo?" as a context for representing and interpreting data in a box plot.

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.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. For example, "How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.
MGSE6.SP.2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
MGSE6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.
MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS FOR MATHEMATICAL PRACTICE

1. Make sense of problems and persevere in solving them. Students will make sense of the data distributions by interpreting the measures of center and variability in the context of the situations they represent.
4. Model with mathematics. They use measures of center and variability and data displays (i.e. box plots) to draw inferences about data sets. Students collect data regarding real-world contexts and create models to display and interpret the data.
6. Attend to precision. Students use appropriate terminology when referring data displays and statistical measures.
8. Look for and express regularity in repeated reasoning. They will explain that each quartile of the box and whisker plot is 25% of the data.

ESSENTIAL QUESTIONS

- How can I describe the center and spread of a set of data?
- How can I use data to compare different groups?
- What conclusions can be drawn from data?

MATERIALS
Coin
1 Waldo picture per group
Stop watch or timer

Where's Waldo?

Part 1: Creating and Describing Box Plots
How fast can you find Waldo? Collect data using the steps below to answer this question.

TEACHER

1. Pair students in groups of 2.
2. Have students decide who will be Student 1 (searcher) and Student 2 (timer).
3. Give each pair a copy of "Where's Waldo?" picture side down. Students are not to touch the paper or turn it over until told.

Comment: See accompanied Waldo photo. You may choose to find your own Waldo image online to print in color for a class set.

4. Show the students a picture of Waldo on the board so they know what to look for.

Comment: Many students have never heard of Where's Waldo. So, show them a picture of Waldo from the internet in color on your board if possible!

STUDENT

1. When the teacher says, "GO", Student 1 (the searcher) is to turn the paper over and find Waldo as quickly as possible. Student 2 (the timer) will time how long it takes for Student 1 to find and point to Waldo. Record your time to the nearest second. Keep times in SECONDS and do not convert to minutes

Comment: Watch out for "false finds". They MUST find Waldo before they say they found him!

TEACHER
Collect data from the class on the board.
Comment: Have the students organize how to collect this data on the board. Allow them to come to the board to write answers if you choose.

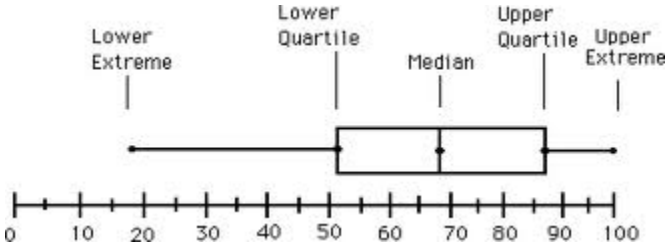
STUDENT

2. Fill in the chart below.

a. Record data	
b. Put the observations in least to greatest order.	
c. Identify the least number of seconds.	
d. Identify the most number of seconds.	
e. Find the median (Q2).	
f. Find the lower quartile (Q1).	
g. Find the upper quartile (Q3).	

3. Create a box plots using the data in the chart.

Sample Answer:



4. What is the attribute being measured in this task?

The time it takes to find Waldo

5. What unit is being used to measure this attribute and why?

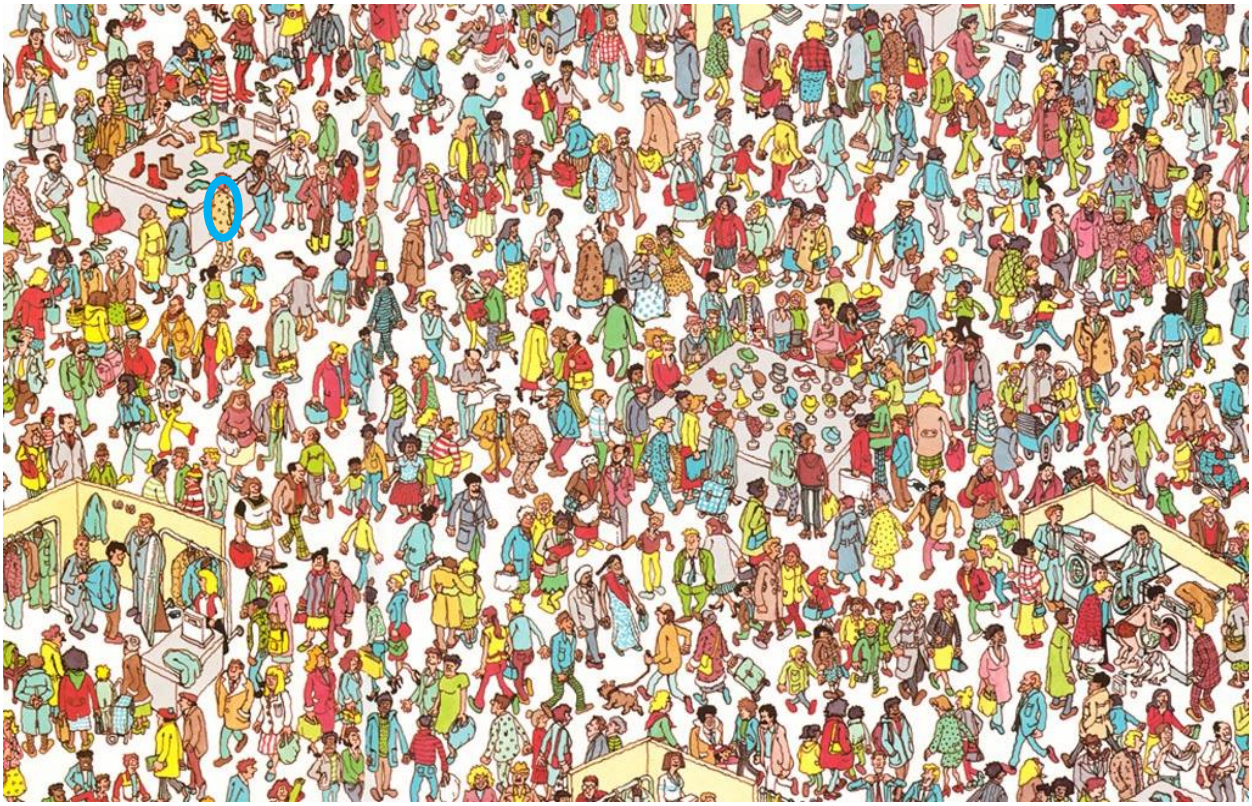
Seconds because they are more precise than minutes

6. Describe the spread of the data for the box plot and explain what this tells you about the data?

Answers will vary depending on specific data.

7. Do you think if we did the same experiment with 30 other random people, we would come up with the same conclusion? Why or why not?

Comment: Teachers may want to discuss what would happen if they collect a whole different set of data from a different set of students. Do they feel that it will yield the same conclusions? Why or why not?
As an extension students may compare data of all classes to determine if their predictions were correct.



Part 2: A Closer Look at Box Plots

Below is a data set of the length of times, in seconds, that it took for nine boys to find Waldo:

7 8 8 10 11 12 13 13 29

1. Write the numbers in order and circle the median.

7 8 8 10 (11) 12 13 13 29

a. About what percent of the values in a data set are below the median? How do you know this?
Approximately 50 %. This is because the median is the middle of the data set, so the median splits the data set in half.

b. About what percent of values in a data set are above the median? How do you know this?
Approximately 50%. Because ½ of all values lie above the median.

2. Find the Upper (Q3) and Lower (Q1) Quartiles of the box plot. Draw a small vertical line where Q1 and Q3 are on the list of numbers above.

Comment: Remember to find the value of Q1 and Q3 in this data set, you will need to find the mean of the two middle values in Q1 and Q3.

7 8 | 8 10 (11) 12 13 | 13 29

The value of Q1 is 8. ***The value of Q3 is 13.***

Emphasize the meaning of the root word in "quartile". The three quartiles (Q1, Q2, and Q3) divide the data into four sections, each containing the same number of data points. Allow students to share examples of familiar words that begin with the root word, "inter-". Discuss the meaning of "inter" and "between". Help students create a definition for the word "interquartile range" by dissecting each part of the term.

a. About what percent of the data distribution are in each quartile?
Approximately 25%

b. About what percent of the values fall ABOVE the lower quartile? ***75%***

c. About what percent of the values fall below the upper quartile? ***75%***

d. The Inter-Quartile Range (IQR) is the size of the "box" of the box plot. The box contains all data between Q1 and Q3. What percent of the data fall between the upper quartile and the lower quartile? ***50%. It is the middle 50% of the data set.***

e. Find the IQR using the values given above. ***13-8=5***

f. Why is the IQR important when using it to describe the data?
It means that 50% of our data set is between 8 and 13. So, the middle 50% of the data is within 5 seconds of the median of 11.

Where's Waldo?

Part 1: Creating and Describing Box Plots

How fast can you find Waldo? Collect data using the steps below to answer this question.

1. Students will work with partners.

2. Students decide who will be Student 1 (searcher) and Student 2 (timer). Students could flip a coin.

3. Give each pair a copy of "Where's Waldo?" picture side down. Students are not to touch the paper or turn it over until told.

4. Look on the board for a picture of Waldo so you will know what to look for.

STUDENT

5. When the teachers says, "GO", Student 1 (the searcher) is to turn the paper over and find Waldo as quickly as possible. Student 2 (the timer) will time how long it takes for Student 1 to find and point to Waldo. Record your time to the nearest second. Keep times in SECONDS and do not convert to minutes

Partner 1 time in seconds_____

Partner 2 time in seconds_____

Using data collected from each group fill in the chart below.

8. Fill in the chart below.

h. Record data	
i. Put the observations in least to greatest order.	
j. Identify the least number of seconds.	
k. Identify the most number of seconds.	
l. Find the median (Q2).	
m. Find the lower quartile (Q1).	
n. Find the upper quartile (Q3).	

9. Create a box plots using the data in the chart.

10. What is the attribute being measured in this task?

11. What unit is being used to measure this attribute and why?

12. Describe the spread of the data for the box plot and explain what this tells you about the data?

13. Do you think if we did the same experiment with 30 other random people, we would come up with the same conclusion? Why or why not?

Part 2: A Closer Look at Box Plots

Below is a data set of the length of times, in seconds, that it took for nine boys to find Waldo:
7 8 8 10 11 12 13 13 29

1.Write the numbers in order and circle the median.

a. About what percent of the values in a data set are below the median? How do you know this?

b. About what percent of values in a data set are above the median? How do you know this?

2. Find the Upper (Q3) and Lower (Q1) Quartiles of the box plot. Draw a small vertical line where Q1 and Q3 are on the list of numbers above.

a. About what percent of the data distribution are in each quartile?

b. About what percent of the values fall ABOVE the lower quartile?

c. About what percent of the values fall below the upper quartile?

d. The Inter-Quartile Range (IQR) is the size of the "box" of the box plot. The box contains all data between Q1 and Q3. What percent of the data fall between the upper quartile and the lower quartile?

e. Find the IQR using the values given above.

f. Why is the IQR important when using it to describe the data?

Guinness Outliers

Back to Task Table

In this task, students will collect a variety of measurement data and then investigate how an extreme outlier can affect the measures of center.

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS FOR MATHEMATICAL PRACTICE

1. Make sense of problems and persevere in solving them. Students will make sense of why measures of center (i.e., mean and median) are sometimes close to one another and sometimes significantly different and how the context of a set of data renders one measure of center more appropriate than the other when describing the data.
3. Construct viable arguments and critique the reasoning of others. Students will explain how measures of center (i.e., mean and median) are affected when an outlier is added to a set of data.
6. Attend to precision. Students must pay careful attention to specified units of measure for each data set. In some cases, students may need to realize that the given unit must be converted.
8. Look for and make use of structure. Students looking for structure will realize that, when adding one value to a data set for which they have already calculated a mean, they can simply add the new value to the sum of the existing values and then divide by one more than the number of existing values in the data set.

MATERIALS

Calculators
Rulers
Tape Measures
Soccer Ball (or similar ball)
Wall Posters for measuring height (cm)
Student Recording Sheet (included)
First 200 digits of Pi Graphic (included)
Pictures of and/or hyperlinks to Guinness World Records (included)

TEACHER COMMENTS

Students often struggle with the task of determining which measure of center is most appropriate to use when summarizing a set of data, particularly in data sets that are not equally distributed around the center. This task provides an engaging learning context for allowing students to make meaning of this phenomenon.

In this task, seven measurement scenarios are presented to students (ideally, in stations around the classroom). Each student will:

1) Juggle a soccer ball as long as they can while being timed by a peer (or teacher)
2) Measure the width of his/her mouth (in cm).
3) Have a peer measure the longest hair on his/her head (in cm).
4) Measure the length of his/her thumb nail.
5) Count his/her total number of fingers and toes.
6) Have a peer measure his/her height (in cm). *A wall poster works really well for this!*
7) Recite as many digits of Pi from memory as possible. *The first 200 digits of Pi are presented in this task – for students who are unfamiliar with this number, allow them to study it for a minute first before attempting to recite.*

In groups of 3 or 4, have students move around the room to visit each station and collect/record data (see student record sheet included in this task). **If students feel uncomfortable performing any of the tasks or measurements, please use professional discretion – you may want to omit one or more scenarios, or possibly even create your own.**

After collecting the data, students should then work to find the group mean and median for each scenario (not for the whole class, but rather, the small group with whom they are working). More than likely, students will not need a calculator for this part of the task.

Have groups share their data with the whole class. Ask the students what they notice about the spread of each data set (e.g., what is the approximate range for each set and why are some ranges much larger than others?).

Now, inform students that they will add an extra data point to their existing sets: the world record holder in each category they just measured. Before sharing the record-holders, ask students to predict what will happen to the measures of center and variability for each data set once the record holders are added? Look for students who say, "They will get bigger," or "The mean and the median will both change dramatically." Do NOT correct these misconceptions – allow the students to experience cognitive dissonance and discover the relationships on their own.

For each data set, share the pictures provided of each world record holder. Instruct students to record this data point as the "Guinness Outlier" in their data sets. After sharing all of the outliers, allow students to use a calculator to find the NEW mean for each data set, as well as the NEW median. When calculating the mean, look to see if students realize that they can simply take the sum of their data and add the outlier and divide by "one more than the number of students in their group." If they don't see this structure (SMP #8), pose the question, "Does anyone know of a way to find the new mean WITHOUT entering all of the data values in the calculator?"

After refiguring the mean and median for each data set, ask the question, "What do you notice?" See if students pick up on the fact that the mean is now significantly higher than the rest of the values in each corresponding data set. If they point this out, be sure to ask them, "Why?" What about the median? Does it also change significantly when an outlier is added to a set of data? Why or why not?

In closing, ask the students to choose which measure of center is most appropriate to use when describing what is "typical" about each data set? If some think the mean is more appropriate than the median, ask them to justify/explain before calling on another student or trying to explain why that is incorrect. At this juncture, most students will better understand that the median is far less misleading when an extreme outlier is present in the data set. If students continue to struggle with this concept, represent the data on an empty number line to show the isolation of the outlier and the clustering of the other values. See if students can use fraction or percent language to explain the spread of the data (e.g., "I see that 4/5 or 80% of the data is clustered between 1cm and 2cm, so that's where most of us were.).

NOTE: Some of the measurement units provided for the outliers in this task are not the same as the units students will use in the classroom portion of the task (e.g., the soccer juggling). See if students pick up on this (SMP #6) rather than telling them that they will need to convert. You might also want to let them figure out how to perform the conversions, rather than simply telling them how to do so.

On December 6, 1995, Nikolai Kutsenko juggled a soccer ball for 24 hours and 30 minutes straight without letting the ball hit the ground!
Source: http://www.recordholders.org/en/list/ball.html
Picture from http://www.20min.ch/2010/img/diashow_html5/1x1.gif

On March 18, 2010, Francisco Domingo Joaquim set the record for widest mouth (unstretched) at 17cm!
Source: http://www.guinnessworldrecords.com/world-records/widest-mouth
Picture from http://www.guinnessworldrecords.com/assets/45549

On May 8, 2004, Xie Qiuping set the record for longest head hair at 563 cm!
Source: http://www.guinnessworldrecords.com/world-records/longest-hair-(female)
Picture from http://www.guinnessworldrecords.com/media/6562234/99247-longest_hair_2.jpg

On September 16, 2013, Christine Walton set the record for longest thumbnail at 107 cm!
Source: http://www.guinnessworldrecords.com/world-records/longest-fingernails-on-a-pair-of-hands-(female)
Picture from http://www.guinnessworldrecords.com/assets/516974

Born on January 9, 1995, Devendra Harne has 25 total fingers and toes. The record was tied on August 10, 2005, when Pranamya Menaria was also born with 25 fingers and toes.
Source: http://www.guinnessworldrecords.com/world-records/most-fingers-and-toes-%28polydactylism%29-living-person
Picture from http://www.enteresan.com/files/img/ent/art/1Rk/CS3/s1M/en-siradisi-10-vucut-uzuv,6,box.jpg

Sultan Kösen, born on December 10, 1982, is the world's tallest living man, at 251 cm!
Source: http://www.guinnessworldrecords.com/world-records/tallest-man-living
Picture from http://presse.niessing.com/presse/SultanFotos/Sultan_Koesen_Niessing_02.jpg

On November 20, 2005, Chao Lu recited 67,890 digits of Pi from memory!
Source: http://www.pi-world-ranking-list.com/index.php?page=lists&category=pi
Picture from http://www.pi-world-ranking-list.com/lists/details/luchao.jpg

First 200 digits of Pi, for students to use with peers as they recite and record the number of digits recalled from memory. Print this and cut out to place in this station.

3.14159265358979323846264338327950288419716939937510
58209749445923078164062862089986280348253421170679
82148086513282306647093844609550582231725359408128 4
81117450284102701938521105559644622948954930381 96

Guinness Outliers (Recording Sheet) Name_____

Directions: *Please complete the measurement activity at each station. For many of the measurement tasks, you will need a partner to help you measure.*

- *Be sure to record your data accurately (attending to the specified units of measure). Record your data in space A; record the data of your group members in spaces B, C, and/or D.*
- *Once you have collected all of your group's data for each measurement task, determine the mean and median of your group. Then, wait for further directions.*

See how many **seconds** you can juggle a soccer ball without letting it hit the floor.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**

Open your mouth (all the way) and have a partner measure the width of your mouth to the nearest **cm**.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**

Have a partner measure the length of the longest hair on your head to the nearest **cm**.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**

Measure the length of your thumb nail to the nearest **cm**.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**
					

Count the total number of **fingers and toes** you have.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**
					

Have a partner measure your height to the nearest **centimeter**.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**
					

From memory, recite as many consecutive **digits** of Pi as you can. Have a partner listen and record your data.

Student				Measure of Center	
A	**B**	**C**	**D**	**Mean**	**Median**
Outlier (world record)				**New Mean**	**New Median**
					

1. What did you notice about the mean and median of your group's data BEFORE including the outlier (world record)?

2. What did you notice about the mean and median of your group's data AFTER including the outlier (world record)?

3. Look at your NEW measures of center for each data set. If someone wanted know what was a "typical" or "average" measure for your group (with the outlier included), which measure of center would you use – mean or median? Explain your thinking (*be sure to cite a specific example or evidence in your explanation).*

Mean, Median, Mode and Range: Formative Assessment Lesson

Back to Task Table

Source: Formative Assessment Lesson Materials from Mathematics Assessment Project
http://map.mathshell.org/materials/download.php?fileid=1360

TASK COMMENTS

Tasks and lessons from the Mathematics Assessment Project are specifically designed to help teachers effectively formatively assess their students. The way the tasks and lessons are designed gives the teacher a clear understanding of what the students are able to do and not do. Within the lesson, teachers will find suggestions and question prompts that will help guide students towards understanding. For more information access the MAP website:
http://www.map.mathshell.org/materials/background.php?subpage=formative

The task, *Mean, Median, Mode and Range*, is a Formative Assessment Lesson (FAL) that can be found at the website:
http://map.mathshell.org/materials/lessons.php?taskid=486&subpage=concept

The FAL document provides a clear lesson design, from the opening of the lesson to the closing of the lesson.

The PDF version of the task can be found at the link below:
http://map.mathshell.org/materials/download.php?fileid=1360

STANDARDS FOR MATHEMATICAL CONTENT

Develop understanding of statistical variability.

MGSE6.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.
MGSE6.SP.2 Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
MGSE6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

STANDARDS FOR MATHEMATICAL PRACTICE

This lesson uses all of the practices with emphasis on:

1. Make sense of problems and persevere in solving them
2. Reason abstractly and quantitatively.
4. Model with mathematics.

ESSENTIAL QUESTIONS

- How can I describe the center for a set of data?
- How can I describe the spread for a set of data?
- How can I use data to compare different groups?
- What conclusions can be drawn from data?
- How can I calculate the mean, median, and range from a frequency chart?

Suzi's Company (Short Cycle Task)

Back to Task Table

Source: Balanced Assessment Materials from Mathematics Assessment Project
http://www.map.mathshell.org/materials/download.php?fileid=1138

TASK COMMENTS

Tasks and lessons from the Mathematics Assessment Project are specifically designed to help teachers effectively formatively assess their students. The way the tasks and lessons are designed gives the teacher a clear understanding of what the students are able to do and not do. Within the lesson, teachers will find suggestions and question prompts that will help guide students towards understanding. For more information access the MAP website:
http://www.map.mathshell.org/materials/background.php?subpage=summative

The task, *Suzi's Company*, is a Mathematics Assessment Project Assessment Task that can be found at the website:
http://www.map.mathshell.org/materials/tasks.php?taskid=383&subpage=apprentice

The PDF version of the task can be found at the link below:
http://www.map.mathshell.org/materials/download.php?fileid=1138

The scoring rubric can be found at the following link:
http://www.map.mathshell.org/materials/download.php?fileid=1139

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS FOR MATHEMATICAL PRACTICE

This task uses all of the practices with emphasis on:

3. Construct viable arguments and critique the reasoning of others.
7. Look for and make use of structure.

Candy Bars (Short Cycle Task)

Back to Task Table

Source: Balanced Assessment Materials from Mathematics Assessment Project
http://www.map.mathshell.org/materials/download.php?fileid=1178

TASK COMMENTS

Tasks and lessons from the Mathematics Assessment Project are specifically designed to help teachers effectively formatively assess their students. The way the tasks and lessons are designed gives the teacher a clear understanding of what the students are able to do and not do. Within the lesson, teachers will find suggestions and question prompts that will help guide students towards understanding. For more information access the MAP website:
http://www.map.mathshell.org/materials/background.php?subpage=summative

The task, *Candy Bars*, is a Mathematics Assessment Project Assessment Task that can be found at the website: http://www.map.mathshell.org/materials/tasks.php?taskid=396&subpage=expert

The PDF version of the task can be found at the link below:
http://www.map.mathshell.org/materials/download.php?fileid=1178

The scoring rubric can be found at the following link:
http://www.map.mathshell.org/materials/download.php?fileid=1179

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS FOR MATHEMATICAL PRACTICE

This task uses all of the practices with emphasis on:

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.
5. Use appropriate tools strategically
6. Attend to precision.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

Culminating Task: Order Up! Fast Food Frenzy

Back to Task Table

In this task, students will analyze fat and calorie content in popular fast food items.

STANDARDS FOR MATHEMATICAL CONTENT

MGSE6.SP.2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
MGSE6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.
MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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

STANDARDS FOR MATHEMATICAL PRACTICE

1. Make sense of problems and persevere in solving them. Students will make sense of the data distributions by interpreting the measures of center and variability in the context of the situations they represent.
2. Reason abstractly and quantitatively. Students reason about the appropriate measures of center or variability to represent a data distribution.
4. Model with mathematics. Students use histograms, box plots and dot plots, to draw inferences about a data set.
6. Attend to precision. Students use appropriate terminology when referring data displays and statistical measures.

ESSENTIAL QUESTIONS

- What is meant by the center of a data set, how is it found and how is it useful when analyzing data?
- How do I create appropriate graphs to represent data?
- How can I describe the spread of a set of data?
- What conclusions can be drawn from data?

Every day in the United States, millions of customers eat fast food. There are so many different types of numerical data and information that we can gather about fast food, and then make decisions based on the evidence found.

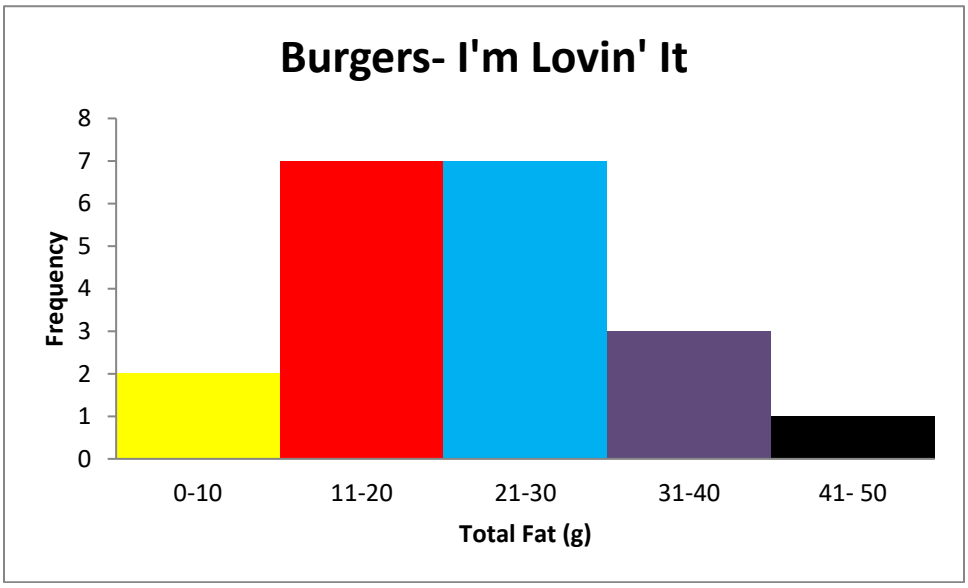
Part 1: Burgers~ I'm Lovin' It!

Sandwiches	Total Fat (g)
Hamburger	9
Cheeseburger	12
Double Cheeseburger	23
McDouble	19
Quarter Pounder®	19
Quarter Pounder® with Cheese	26
Double Quarter Pounder® with Cheese	42
Big Mac®	29
Big N' Tasty®	24
Big N' Tasty® with Cheese	28
Angus Bacon & Cheese	39
Angus Deluxe	39
Angus Mushroom & Swiss	40
Filet-O-Fish®	18
McChicken ®	16
McRib ®†	26
Premium Grilled Chicken Classic Sandwich	10
Premium Crispy Chicken Classic Sandwich	20
Premium Grilled Chicken Club Sandwich	17
Premium Crispy Chicken Club Sandwich	28

Source: McDonald's USA

1. Write 3 statistical and 3 non-statistical questions about the data in the previous table.

2. Create a histogram from the data above. Into which interval do the majority of the sandwiches fall?



bin	*Frequency*
0-10	2
11-20	7
21-30	7
31-40	3
41- 50	1
More	0

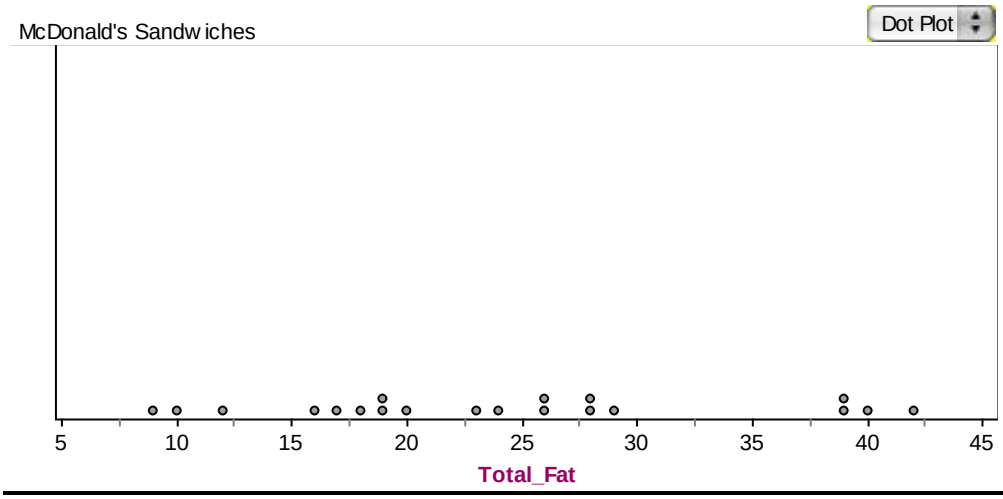
3. What is the total number of observation (pieces of data) that we are analyzing? n=_____
n=20

Measures of Center

4. What is the mean (average) number of calories in a sandwich from McDonalds?
The mean of fat grams is 24.2.

5. In context of the data given, describe the meaning of the mean:
The mean number of fat grams that can be found in a McDonalds Sandwich is 24.2 grams.

6. Create a dot plot of the data. Be sure to make your tic marks evenly spaced on the axis below:



Number of Fat Grams

Measures of Spread

7. What is the highest amount of fat grams? _____ What is the least amount of fat grams?_____
The sandwich with the highest amount of fat grams is the Double Quarter Pounder with Cheese with 42 fat grams. The sandwich with the least amount of fat grams is the Hamburger with 9 fat grams.

8. What is the range of the data?_____
The range of the data is the highest value minus the lowest value. $42 - 9 = 33$

9. Explain what the range of the data means in terms of fat grams.
The range of fat grams that can be found in McDonald's sandwiches is 33 grams. This helps us conclude that there is a wide variety in the number of fat grams found in McDonald's sandwiches.

Shape

10. Are there about the same amount of data on both sides of the mean? Look at your dot plot to inspect this (this is called a "symmetric" shape). If not, then your dot plot might be skewed (the bulk of the data on one side, while the other side tails off).
Looking at both sides of the mean on the dot plot, there appears to be approximately the same number of observations on each side of the mean. So this data set is approximately symmetric

11. Critical Thinking: According the USDA, 30% of your daily calorie intake should come from fat. For a person that consumes 2,000 calories/day, this is 67 grams of fat.

Based on this information, which type of sandwich would you NOT suggest an individual to eat and why?

Sample Answer:
The double quarter pounder with cheese is clearly the highest in fat, with 42 fat grams. I wouldn't eat any of the sandwiches above 29 fat grams. Since this is only one meal of the day, it would be difficult to stay within the USDA recommendation when eating almost half of your fat grams in one meal.

Part 2: Biggie Size those Fries…. Or Maybe Not!!!

	Small	Medium	Large
McDonalds	230	380	500
Wendy's	330	410	540
Burger King	340	440	540
Dairy Queen		310	500
Sonic	204	326	448
Steak and Shake	240	440	640
Chick-fil-A	290	380	430
Arby's	410	540	640

1. Write 3 statistical and 3 non-statistical questions about the data in the previous table.

2. Find the **5 Number Summary** for each size of fries:

5 Number Summary	Small Fry	Medium Fry	Large Fry
Minimum	*204*	*310*	*430*
Quartile 1	*230*	*353*	*474*
Median (Quartile 2)	*290*	*395*	*520*
Quartile 3	*340*	*440*	*590*
Maximum	*410*	*540*	*640*

3. Make a box plot for each size of fries. All three box plots should use the same intervals. The horizontal axis should be marked with amount of calories.

Create the three box plots here.

Students will create a box plot for each size of French fries. They will have one for small, medium and large fries.

Spread

4. What is the range of calories for small, medium, and large fries?
Small=_____ Medium=_____ Large=_____

Comment
It is useful for students to know how the data is spread out.

Solution
The range of calories is 136 calories for small, 230 calories for medium, and 210 calories for a large fry.

5. Using the five-number summary, find the Inter-Quartile Range (IQR) for small, medium, and large fries.
Small Fry IQR=_____ Medium Fry IQR=_____ Large Fry IQR=_____
Solution: Small Fry IQR: 340-230=110. Medium Fry IQR: 440-353=87. Large Fry IQR: 590-474=116

6. Which size fry has the greatest range? _____ Which size fry has the greatest IQR? _____.
Solution: Medium fries have the greatest range and large fries have the greatest IQR.

7. In trying to determine which size fry has the more consistent amount of calories, would you compare the ranges or the IQR? Explain why you chose the measure you did.
Solution: The IQR is the middle 50% of the data, so it may be more useful to look at instead of the range.

Shape

8. Describe the shape of each box plot? Justify your response.
Solution: No, none of these boxes have symmetry about the median.

9. Overall, which restaurant had fries that were very high in calories?_____
Solution: Arby's

10. Which restaurant had fries that were overall lowest in calories? _____
Solution: Sonic

Variability

11. Name some factors that might cause one restaurant's large fry order be much higher in calories than another restaurant's large fry. These factors are called *uncontrolled variables.*
Sample Solution: Each restaurants' Small, Medium, and Large fry may vary in size. The type of oil that the restaurant uses to fry can make calories higher or lower.

Part 3: I'm Thirsty!
Below is a list of drinks to choose from at McDonalds.

	Calories in 21 fluid ounces	Grams of Sugar in 21 fluid ounces
Coca-Cola Classic®	210	40
Dasani Water®	0	0
Diet Coke®	0	0
Sprite®	210	56
Hi-C® Orange	240	64
Orange Juice (22 oz)	280	58
Powerade ®	150	31
Sweet Tea	180	45
Iced Mocha (22 oz)	340	41
Hot Chocolate (20 oz)	460	54
Frappe Mocha (22 oz)	680	87
Strawberry Bananna Smoothie (22 oz)	330	70
Chocolate MGSEafe Shake (22 fl oz)	720	98

1. Create a histogram displaying the number of calories found in the popular drinks at McDonalds.
Comment: Histograms can look very different, if different class widths are chosen in the frequency table. (Example counting from 0-10 for one histogram bar versus counting from 0-20). So, histograms can look different. You can suggest a class width size to the class

2. Find the mean and the median of the calorie data. Which measure of center would best describe the caloric data and why?
Solution: The mean is 292.31 and the median is 240 calories. The median is a better choice for describing data because it is the center of the data. The mean is strongly affected by outlier (such as 0 calories and 720 calories in this case).

3. Create a histogram in the graphing calculator using the grams of sugar data. Sketch a graph of your histogram below.
Comment: Histograms can look very different, if different class widths are chosen in the frequency table. (Example counting from 0-10 for one histogram bar versus counting from 0-20). So, histograms can look different. You can suggest a class width size to the class.

4. Find your mean and median using the 1-VAR STATS tool under the STATS menu. Which would be best in describing grams of sugar in drink at McDonalds?
Solution: Mean is 49.54 grams and the median is 54 grams. The median is the best description since it is not influenced by outliers.

CULMINATING TASK: ORDER UP! FAST FOOD FRENZY

Every day in the United States, millions of customers eat fast food. There are so many different types of numerical data and information that we can gather about fast food, and then make decisions based on the evidence found.

Part 1: Burgers~ I'm Lovin' It!

Sandwiches	Total Fat (g)
Hamburger	9
Cheeseburger	12
Double Cheeseburger	23
McDouble	19
Quarter Pounder®	19
Quarter Pounder® with Cheese	26
Double Quarter Pounder® with Cheese	42
Big Mac®	29
Big N' Tasty®	24
Big N' Tasty® with Cheese	28
Angus Bacon & Cheese	39
Angus Deluxe	39
Angus Mushroom & Swiss	40
Filet-O-Fish®	18
McChicken ®	16
McRib ®†	26
Premium Grilled Chicken Classic Sandwich	10
Premium Crispy Chicken Classic Sandwich	20
Premium Grilled Chicken Club Sandwich	17
Premium Crispy Chicken Club Sandwich	28

Source: McDonald's USA

1. Write 3 statistical and 3 non-statistical questions about the data in the previous table.

2. Create a histogram from the data above. Into which interval do the majority of the sandwiches fall?

3. What is the total number of observation (pieces of data) that we are analyzing? n=_____

Measures of Center

4. What is the mean (average) number of calories in a sandwich from McDonalds?

5. In context of the data given, describe the meaning of the mean:

6. Create a dot plot of the data. Be sure to make your tic marks evenly spaced on the axis below:

Number of Fat Grams

Measures of Spread

7. What is the highest amount of fat grams? _____ What is the least amount of fat grams?_____

8. What is the range of the data?_____

9. Explain what the range of the data means in terms of fat grams.

Shape

11. Are there about the same amount of data on both sides of the mean? Look at your dot plot to inspect this (this is called a "symmetric" shape). If not, then your dot plot might be skewed (the bulk of the data on one side, while the other side tails off).

12. Critical Thinking: According the USDA, the 30% of your daily calorie intake should come from fat. For a person that consumes 2,000 calories/day, this is 67 grams of fat.

Based on this information, which type of sandwich would you NOT suggest an individual to eat and why?

Part 2: Biggie Size those Fries…. Or Maybe Not!!!

	Small	Medium	Large
McDonalds	230	380	500
Wendy's	330	410	540
Burger King	340	440	540
Dairy Queen		310	500
Sonic	204	326	448
Steak and Shake	240	440	640
Chick-fil-A	290	380	430
Arby's	410	540	640

1. Write 3 statistical and 3 non-statistical questions about the data in the previous table.

2. Find the 5 Number Summary for each size of fries:

5 Number Summary	Small Fry	Medium Fry	Large Fry
Minimum			
Quartile 1			
Median (Quartile 2)			
Quartile 3			
Maximum			

3. Make a box plot for each size French fry using the 5 number summary of data above. All three box plots should use the same intervals. The horizontal axis should be marked with amount of calories.

Spread

4. What is the range of calories for small, medium, and large fries?

Small=_____ Medium=_____ Large=_____

5. Using the five-number summary, find the Inter-Quartile Range (IQR) for small, medium, and large fries.

Small Fry IQR=_____ Medium Fry IQR=_____ Large Fry IQR=_____

6. Which size fry has the greatest range? _____ Which size fry has the greatest IQR? _____.

7. In trying to determine which size fry has the more consistent amount of calories, would you compare the ranges or the IQR? Explain why you chose the measure you did.

Shape

8. Do the box plots appear to be symmetrical? Justify your response

9. What is the best measure of center for each size of French Fry and why?

10. Overall, which restaurant had fries that were very high in calories?_____

11. Which restaurant had fries that were overall lowest in calories? _____

Variability

12. Name some factors that might cause one restaurant's large fry order be much higher in calories than another restaurant's large fry. These factors are called *uncontrolled variables.*

Part 3: I'm Thirsty!

Below is a list of drinks to choose from at McDonalds.

	Calories in 21 fluid ounces	Grams of Sugar in 21 fluid ounces
Coca-Cola Classic®	210	40
Dasani Water®	0	0
Diet Coke®	0	0
Sprite®	210	56
Hi-C® Orange	240	64
Orange Juice (22 oz)	280	58
Powerade ®	150	31
Sweet Tea	180	45
Iced Mocha (22 oz)	340	41
Hot Chocolate (20 oz)	460	54
Frappe Mocha (22 oz)	680	87
Strawberry Bananna Smoothie (22 oz)	330	70
Chocolate MGSEafe Shake (22 fl oz)	720	98

1. Create a histogram displaying the number of calories found in the popular drinks at McDonalds.

2. Find the mean and the median of the calorie data. Which measure of center would best describe the caloric data and why?

3. Create a histogram in the graphing calculator using the grams of sugar data. Sketch a graph of your histogram below.

4. Find your mean and median using the 1-VAR STATS tool under the STATS menu. Which would be best in describing grams of sugar in drink at McDonalds?

Unit Web Links

Back to Task Table

Develop understanding of statistical variability.
http://illuminations.nctm.org/Search.aspx?view=search&cc=2059_2107

MGSE6.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. For example, "How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.

https://www.illustrativemathematics.org/content-standards/6/SP/A/1/tasks/703
https://www.illustrativemathematics.org/content-standards/6/SP/A/1/tasks/1040
https://www.illustrativemathematics.org/content-standards/6/SP/A/1/tasks/2008

MGSE6.SP.2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.

https://www.illustrativemathematics.org/content-standards/6/SP/A/2/tasks/1199
https://www.illustrativemathematics.org/content-standards/6/SP/A/2/tasks/2043
https://www.illustrativemathematics.org/content-standards/6/SP/A/2/tasks/1026
https://www.illustrativemathematics.org/content-standards/6/SP/A/2/tasks/2100

MGSE6.SP.3 Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

https://www.illustrativemathematics.org/content-standards/6/SP/A/3/tasks/2097
http://www.learner.org/courses/learningmath/data/session5/part_c/balancing.html
http://www.shodor.org/interactivate/lessons/IntroStatistics/
http://www.shodor.org/interactivate/activities/PlopIt/

Summarize and describe distributions.

MGSE6.SP.4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

https://www.illustrativemathematics.org/content-standards/6/SP/B/4/tasks/877
https://www.illustrativemathematics.org/content-standards/6/SP/B/4/tasks/2043
https://www.illustrativemathematics.org/content-standards/6/SP/B/4/tasks/2047
https://www.illustrativemathematics.org/content-standards/6/SP/B/4/tasks/1026
https://www.illustrativemathematics.org/content-standards/6/SP/B/4/tasks/2100
http://www.learner.org/courses/learningmath/data/session4/part_d/drawing.html
http://illuminations.nctm.org/Activity.aspx?id=3476
http://illuminations.nctm.org/lesson.aspx?id=1688

MGSE6.SP.5 Summarize numerical data sets in relation to their context, such as by:

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).

https://www.illustrativemathematics.org/content-standards/6/SP/B/5/tasks/877
https://www.illustrativemathematics.org/content-standards/6/SP/B/5/tasks/2043
https://www.illustrativemathematics.org/content-standards/6/SP/B/5/tasks/2047
https://www.illustrativemathematics.org/content-standards/6/SP/B/5/tasks/2054
http://www.openmiddle.com/lower-and-upper-quartiles-with-constraints/
http://www.openmiddle.com/median-with-constraints/
http://www.openmiddle.com/mean-median-and-range/

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

https://www.illustrativemathematics.org/content-standards/6/SP/B/5/tasks/1199
https://www.illustrativemathematics.org/content-standards/6/SP/B/5/tasks/2048