

## Table of Contents

<b>CHAPTER 4: STATISTICS (6 WEEKS)</b>	<b>2</b>
4.0 Anchor Problem: The Electoral College	8
SECTION 4.1: MEASURES OF SHAPE	11
4.1a Class Activity: What is a Statistical Question?	12
4.1a Homework: What is a Statistical Question?	17
4.1b Class Activity: Creating and Analyzing Dot Plots	20
4.1b Homework: Creating and Analyzing Dot Plots	27
4.1c Class Activity: Creating and Analyzing Histograms	31
4.1c Homework: Creating and Analyzing Histograms	37
4.1d Self Assessment: Section 3.1	42
SECTION 4.2: MEASURES OF CENTER	47
4.2a Class Activity: How Much is a Handful?	48
4.2a Homework: Mean as a Fair Share	51
4.2b Class Activity: Finding the Mean	52
4.2b Homework: Finding the Mean	57
4.2c Class Activity: Mean as a Balance Point	60
4.2c Homework: Mean as a Balancing Point	66
4.2d Class Activity: Median and Mode	69
4.2e Class Activity: Choosing the Best Measure of Center	73
4.2e Homework: Choosing the Best Measure of Center	79
4.2f Self Assessment: Section 3.2	82
SECTION 4.3: MEASURES OF VARIABILITY	85
4.3a Class Activity: Variability and Mean Absolute Deviation	86
4.3a Homework: Variability and Mean Absolute Deviation	96
4.3b Class Activity: The Interquartile Range	102
4.3b Homework: The Interquartile Range	108
4.3c Class Activity: Box Plots	112
4.3c Homework: Box Plots	118
4.3d Class Activity: Analyzing Box Plots	122
4.3d Homework: Analyzing Box Plots	127
4.3e Self Assessment: Section 3.3	128
SECTION 4.4: INTERPRET AND DRAW CONCLUSIONS ABOUT DATA	132
4.4a Class Activity: Make, Interpret and Draw Conclusions About Data Part 1	133
4.4a Homework: Make, Interpret, and Draw Conclusions About Data Part 1	136
4.4b Class Activity: Make, Interpret, and Draw Conclusions About Data Part 2	138
4.4b Homework: Make, Interpret, and Draw Conclusions About Data Part 2	144
4.4c Self Assessment: Section 3.4	146

# Chapter 4: Statistics (6 Weeks)

## Common Core Standard(s)

- 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.* (6.SP.1)
- 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. (6.SP.2)
- 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. (6.SP.3)
- Display numerical data in plots on a number line, including dot plots, histograms, and box plots. (6.SP.4)
- Summarize numerical data sets in relation to their context, such as by: (6.SP.5)
  - a) Reporting the number of observations.
  - b) Describing the nature of the attribute under investigation, including how it was measured and its unit of measurements.
  - c) Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
  - d) Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

**Vocabulary:** statistical question, numerical data, categorical data, data distribution, attribute, dot plot, frequency table, histogram, shape, cluster, gap, peak, skew, symmetric, outlier, measures of center, mean, absolute deviation, median, mode, measures of variability, spread, mean absolute deviation (MAD), maximum (max), minimum value (min), range, quartiles, 1<sup>st</sup> quartile (lower quartile), 3<sup>rd</sup> quartile (upper quartile), the 5-number summary, interquartile range (IQR)

## Chapter Overview:


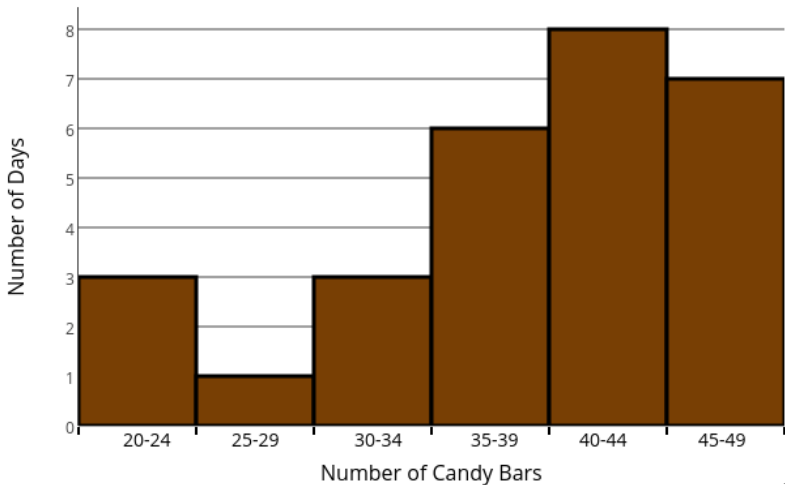

In this chapter students build on their knowledge and experience in data analysis developed in previous grades. They begin the chapter by determining “What is a statistical question?” They learn that statistical questions involve variability in the data collected to answer them. Investigating statistical questions helps students begin to develop an understanding of statistical variability. As students continue to expand their knowledge in data analysis they begin to characterize data distributions by measures of shape, center, and spread. These measures guide the investigation of data analysis throughout the chapter. Working with counts or measurements, students display data with dot plots, histograms, or box plots in order to analyze the shape of the data. They use terms such as cluster, peak, gap, symmetric, skew, and outlier to describe the shape of the data in these graphical displays. In 6<sup>th</sup> grade, students develop well-defined numerical measures that describe the center and spread. The center is measured by the mean, median or sometimes mode. The variability or spread is measured by the mean absolute deviation (MAD), range, or interquartile range (IQR). Once students become more familiar with the components of shape, center, and spread they use them together to summarize and describe distributions, often comparing two or more data sets. They take into account the context in which the data is presented and answer statistical questions.


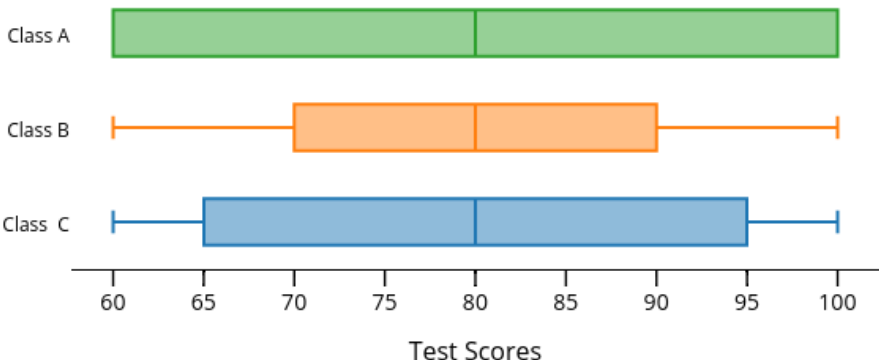

## **Connections to Content:**

Prior Knowledge: In previous grades, students use their work with data to build the foundation for statistical study in 6<sup>th</sup> grade and beyond and to strengthen and apply their study of arithmetic. As early as kindergarten, students begin to classify and sort objects into categories by count. Throughout elementary school they organize, represent, and interpret categorical data and measurement data (See Table 1 in the K-3, Categorical Data; Grades 2-5, Measurement Data Progressions document). Their work with data is closely related to the number line and fractions. They ask and answer questions about the total number of data points within each category and make comparisons. They display their data on line plots (essentially the same thing as dot plots), picture graphs, and bar graphs and answer questions about data given in these plots as well. This investigative process is continued and extended in a more formal statistical approach in 6<sup>th</sup> grade.

Future Knowledge: In Grade 7, students move from concentrating on analysis of data to production of data. They understand that good statistical answers depend on a well-developed plan for collecting data. They investigate random sampling, and in turn, concepts related to probability. Until 8<sup>th</sup> grade, almost all encounters with data analysis have been with univariate data (collections of counts with one variable or characteristic). In 8<sup>th</sup> grade, students extend their knowledge of shape, center, and spread to the analysis of bivariate data, (collection of counts with 2 variables or characteristics) as related to their work with linear functions. In high school statistics, students enhance their understanding of data analysis to give more precise answers to deeper statistical questions. For example, they use standard deviation as a measure of variability that is based on many of the same principals as the mean absolute deviation.

# MATHEMATICAL PRACTICE STANDARDS

	<p><b>Make sense of problems and persevere in solving them.</b></p>	<p>Roman thinks that his school needs to get another vending machine in the school cafeteria because there is always a long line to use the machine at lunch time. The principal has told everyone that they cannot get another vending machine unless they can show that on average at least 40 candy bars are sold each day. Roman conducts a survey and recorded his results in the histogram below.</p> <p style="text-align: center;">Daily Number of Candy Bars Sold from the Vending Machines</p>  <table><caption>Data for Histogram: Daily Number of Candy Bars Sold</caption><tr><th>Number of Candy Bars</th><th>Number of Days</th></tr><tr><td>20-24</td><td>3</td></tr><tr><td>25-29</td><td>1</td></tr><tr><td>30-34</td><td>3</td></tr><tr><td>35-39</td><td>6</td></tr><tr><td>40-44</td><td>8</td></tr><tr><td>45-49</td><td>7</td></tr></table> <p>What possible arguments could Roman give to his principal to convince him that on average at least 40 candy bars are sold from the vending machine each day?</p> <p><i>The example problem given above shows how students must make sense of practical problems and turn them into statistical investigations. They must make sense of what statistical arguments can be made about the data. They must determine what statistical measures might be used to support their arguments and how to go about finding them. Throughout the solving process, students must stop and evaluate their progress. Once they have arrived at a desired solution they can look back and evaluate if their outcome makes sense given the context in which the data is presented.</i></p>	Number of Candy Bars	Number of Days	20-24	3	25-29	1	30-34	3	35-39	6	40-44	8	45-49	7
Number of Candy Bars	Number of Days															
20-24	3															
25-29	1															
30-34	3															
35-39	6															
40-44	8															
45-49	7															
	<p><b>Reason abstractly and quantitatively.</b></p>	<p>A family of five goes out to lunch. The numbers below show the cost of each item that a person ordered for lunch. They decide to split the bill so that everyone pays the same amount. What is the average cost per person for dinner?</p> <p style="text-align: center;">\$8.25      \$8.70      \$7.50      \$6.95      \$9.35</p> <p>Discuss how to find the mean for any given set of data.</p> <p><i>As you discuss how to find the mean for any given set of data students are abstracting the solving process. They must understand that for any data set they can sum the data and divide by the number of data values to get the mean. They come to this understanding by reasoning through the problems and repeatedly creating “fair shares”.</i></p>														

	<p><b>Construct viable arguments and critique the reasoning of others.</b></p>	<p>The three box plots below represent the test scores for three different classes. Examine each plot and then discuss the questions that follow.</p>  <p style="text-align: center;">Test Scores</p> <ol style="list-style-type: none"> <li>What is the same about these box plots and what is different?</li> <li>Find the IQR for each plot and use them to compare the variability of each set of class scores.</li> <li>Make an argument for each class that supports the claim that this class performed the best on the test.</li> </ol> <p><i>The study of statistical analysis requires students to construct arguments that are based off of statistical measures. As students compose these arguments they use stated assumptions, definitions, and previously established results. In many instances the interpretation and analysis of statistical information is subjective and well-constructed and viable arguments can be made for many cases. In the example above, students must make an argument that supports the success of each class's test score data. They can see that such arguments depend on preference and the interpretation of success for this test.</i></p>
	<p><b>Model with mathematics.</b></p>	<p>Below are the ages of 20 people that sing in a choir together.</p> <p>16, 21, 24, 29, 30, 30, 32, 32, 32, 34, 34, 35, 39, 42, 42, 43, 45, 47, 52, 61</p> <ol style="list-style-type: none"> <li>Describe the distribution of ages for people in this choir. Be sure to describe shape, center and variability.</li> <li>What is the (arithmetic) average age for people in this choir? State what kind of numerical measure you used to answer this question, explain your choice.</li> <li>Are most of the people in the choir around the same age? State what kind of numerical measure you used to answer this question, explain your choice.</li> </ol> <p><i>Modeling with mathematics is a major theme throughout this chapter. Not only do students need to know how to interpret data that is displayed graphically but they also need to be able to construct their own models of data. They use dot plots, histograms, and box plots to analyze and draw conclusions about the distribution of data in a variety of contexts.</i></p>

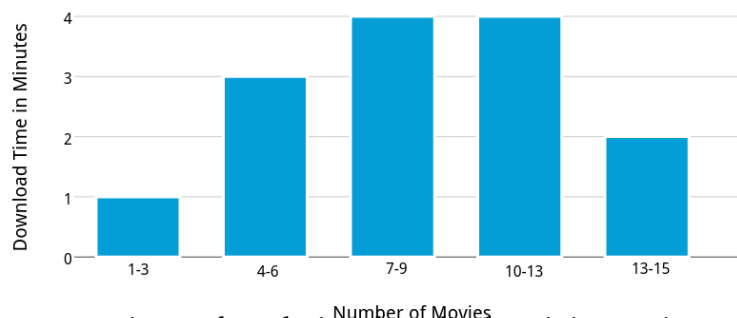


**Attend to precision.**

The table given below shows the time that it takes to download different movies.

Time to download a movie (minutes)	1 – 3	4 – 6	7 – 9	10 – 12	13 – 15
Number of movies	1	3	4	4	2

Carmen made a histogram of the data and has made many mistakes. Find at least two of her mistakes, fix them on the histogram and justify why your reasoning is correct. **Movie Download Times**



*When constructing and analyzing data displays it is very important to pay attention to the scale as this can greatly affect how the data is viewed. This is especially important when students interpret the shape of the data or compare two or more plots. As students attend to precision in constructing data displays they will be able to clarify the correspondence that is occurring within that data distribution that data distribution.*





**Look for and make use of structure.**

The most recent test scores for Mr. Petrov's science class are shown in the table.

Mr. Petrov's Test Scores					
79	91	72	89	84	77
71	82	80	76	91	83
80	79	82	81	78	78
82	72	86	81	80	64
83	79	82	84	81	81
92	80	80	87	83	81

- Make a dot plot of the data. Be sure to label your number line and give it a title.
- How many students are in Mr. Petrov's science class?
- Which test score was earned by the largest number of students?
- Describe any peaks, clusters, or gaps in the data by marking them on the plot.
- What is the overall shape of the data? Justify your answer.
- Mr. Petrov asks, "What is the most typical score for this test?" Use the distribution of data to answer this question.

*Students use structure when analyzing the shape of a data distribution. While identifying peaks, clusters, gaps, skewness, symmetry, and outliers students infer more knowledge about the characteristics of the data. While looking at shape is not numerical structure it is graphical structure and being able to identify what that structure infers about the data is imperative in its analysis.*

	<p>Use appropriate tools strategically.</p>	<p>The high scores for a popular video game are shown below. Li wants to know what the typical high score for this video game is.</p> <table><tr><td>90</td><td>78</td><td>62</td><td>84</td><td>92</td><td>66</td><td>77</td><td>6</td><td>65</td><td>53</td></tr><tr><td>79</td><td>71</td><td>71</td><td>59</td><td>51</td><td>69</td><td>21</td><td>35</td><td>44</td><td>52</td></tr><tr><td>27</td><td>38</td><td>35</td><td>50</td><td>105</td><td>71</td><td>60</td><td>48</td><td>57</td><td>42</td></tr><tr><td>22</td><td>68</td><td>56</td><td>45</td><td>41</td><td>30</td><td>64</td><td>63</td><td>65</td><td>60</td></tr></table> <p>Li begins to make a dot plot for this data but soon realizes that a dot plot is not the best graphical display for this data. Discuss why a dot plot may not be best for this data and see if you can come up with a better way to represent the data.</p> <p><i>Throughout this chapter students will use graphical displays to interpret the shape, center, and spread of the data. These graphical displays are mathematical tools. It is important to choose a graphical display that best suits your data, and your desired measurement. In the example above, the range of the data and the sample space is large, so a dot plot may not be as desirable as a histogram where you can use intervals that summarize the values. At other times a box plot might be preferred because this shows a “snap shot” of the 5-number summary and a visual picture of the variability within the data distributions. Students must also choose appropriate numerical measures of center (mean, median, mode) and variability (MAD, Range, IQR) when interpreting data.</i></p>	90	78	62	84	92	66	77	6	65	53	79	71	71	59	51	69	21	35	44	52	27	38	35	50	105	71	60	48	57	42	22	68	56	45	41	30	64	63	65	60
90	78	62	84	92	66	77	6	65	53																																	
79	71	71	59	51	69	21	35	44	52																																	
27	38	35	50	105	71	60	48	57	42																																	
22	68	56	45	41	30	64	63	65	60																																	
	<p>Look for and express regularity in repeated reasoning.</p>	<p>What does the shape of the data distribution tell you about which measure of center to use to summarize the data?</p> <p><i>After students have analyzed the shape and center of a series of data distributions they begin to understand through repeated reasoning that generally if the shape of the data is fairly symmetrical then the mean is a good measure of center. If the shape of the data is skewed or there are significant outliers, then a good measure of center is the median. Similarly, it is through repeatedly analyzing the shape, center, and spread of several data distributions that they begin to understand how to interpret data, make meaningful conclusions, and answer statistical questions.</i></p>																																								



## 4.0 Anchor Problem: The Electoral College

(Illustrative Mathematics Task)



Unlike many elections for public office where a person is elected strictly based on the results of a popular vote (i.e., the candidate who earns the most votes in the election wins), in the United States, the election for President of the United States is determined by a process called the Electoral College. According to the National Archives, the process was established in the United States Constitution "as a compromise between election of the President by a vote in Congress and election of the President by a popular vote of qualified citizens."

(<http://www.archives.gov/federal-register/electoral-college/about.html> accessed September 4, 2012).

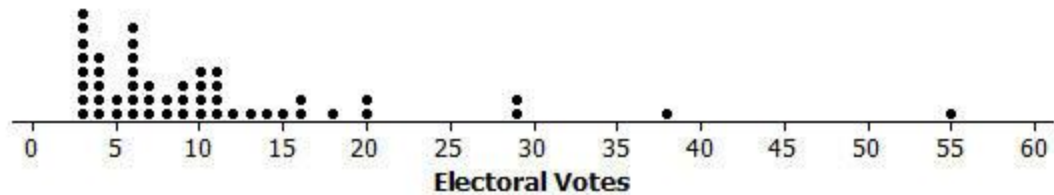
Each state receives an allocation of electoral votes in the process, and this allocation is determined by the number of members in the state's delegation to the US Congress. This number is the sum of the number of US Senators that represent the state (always 2, per the Constitution) and the number of Representatives that the state has in the US House of Representatives (a number that is directly related to the state's population of qualified citizens as determined by the US Census). **Therefore the larger a state's population of qualified citizens, the more electoral votes it has.** Note: the District of Columbia (which is not a state) is granted 3 electoral votes in the process through the 23rd Amendment to the Constitution.

The following table shows the allocation of electoral votes for each state and the District of Columbia for the 2012, 2016, and 2020 presidential elections. (<http://www.archives.gov/federal-register/electoral-college/allocation.html> accessed September 4, 2012).

State	Electoral Votes	State	Electoral Votes	State	Electoral Votes
Alabama	9	Kentucky	8	North Dakota	3
Alaska	3	Louisiana	8	Ohio	18
Arizona	11	Maine	4	Oklahoma	7
Arkansas	6	Maryland	10	Oregon	7
California	55	Massachusetts	11	Pennsylvania	20
Colorado	9	Michigan	16	Rhode Island	4
Connecticut	7	Minnesota	10	South Carolina	9
Delaware	3	Mississippi	6	South Dakota	3
District of Columbia	3	Missouri	10	Tennessee	11
Florida	29	Montana	3	Texas	38
Georgia	16	Nebraska	5	Utah	6
Hawaii	4	Nevada	6	Vermont	3
Idaho	4	New Hampshire	4	Virginia	13
Illinois	20	New Jersey	14	Washington	12
Indiana	11	New Mexico	5	West Virginia	5
Iowa	6	New York	29	Wisconsin	10
Kansas	6	North Carolina	15	Wyoming	3



1. Which state has the most electoral votes? How many votes does it have?
2. Based on the given information, which state has the second highest population of qualified citizens?
3. Here is a dot plot of the distribution.



What is the shape of this distribution? Imagine that someone you are speaking with is unfamiliar with these shape terms. Describe clearly and in the context of this data set what the shape description you have chosen means in terms of the distribution.

4. Does the dot plot lead you to think that any states are outliers in terms of their number of electoral votes? Explain your reasoning, and if you do believe that there are outlier values, identify the corresponding states.
5. What numerical measure or number would you use for describing the “middle” of this data set? Why did you choose this number?

6. Determine the value of the median for this data set (electoral votes).
7. Find the 5-Number Summary values for this data set.
8. Make a box plot for this data set. Use the plot and the IQR to discuss the variability of the data set.

## Section 4.1: Measures of Shape

### Section Overview:

Students begin this section by learning that a statistical question is a question that anticipates variability in the data related to the question and accounts for it in its answers. It is the variability of the data that we try to summarize in statistical analysis. In this section, students will focus on interpreting the shape of the data. To do so they interpret and construct dot plots and histograms of the data. They analyze key features of these graphical displays by looking for clusters, peaks, gaps, and outliers. They also look for whether the data appears to be skewed or symmetric. They answer questions about the data from the data displays including questions that relate to the number observations, the nature of the attribute under investigation, and the context in which the data was gathered.

### Concepts and Skills to Master in this Section:

*By the end of this section, students should be able to:*

1. Recognize a statistical question as one that anticipates variability in the data.
2. Identify if the set of data used to answer a statistical question is numerical or categorical.
3. Display numerical data in a dot plot, use key terms to describe its shape, and interpret the distribution of data.
4. Display numerical data in a histogram, use key terms to describe its shape, and interpret the distribution of data.
5. Summarize a numerical data set in relation to its context by reporting on the number of observations, identifying how the data was measured, and its units of measurement.

#### 4.1a Class Activity: What is a Statistical Question?

Jada, Marco, Ben, Audrey, and Beau each got a different Lego set for Christmas. They each dump out all of their pieces onto the floor to get a better look at what kind of Legos their set includes. As they each sort through their Legos they begin to think about some questions. Their questions are listed below.



#### Lego Questions

1. How many Legos are in each set?
2. How many red Legos are in each set?
3. How much time is needed for each person to complete the Lego set?
4. How much did Marco's Lego set cost?
5. How many Lego pieces have 4 studs (circular bumps on top of the Lego) on them in each set?

You can use statistics to answer most of these questions. **Statistics** is the science of collecting, analyzing, and interpreting data that answer questions. In order to use statistics to answer questions the question posed must be a statistical question.

**A statistical question is question that generates a variety of answers rather than a single answer. This means that the data related to the question varies or has *variability*.**

Part 1: Identify which of the questions on the previous page are statistical questions. For each question, explain why it is or is not a statistical question.







Part 2: Add two more statistical questions related to the Lego sets.

6.

7.

**Directions:** Given each pair of questions below, circle the question that is a statistical question and then explain why.

<p>8.</p> <p>a. How many siblings do you have? </p> <p>b. How many siblings does each student in your class have?</p>	<p>9.</p> <p>a. How many homes were sold each day this month? </p> <p>b. How many homes were sold this month?</p>
<p>10.</p> <p>a. On a scale from 1 to 10 how much pain do you feel when you get a shot at the doctor's office? </p> <p>b. On a scale from 1 to 10 how much pain do your classmates feel when they get a shot at the doctor's office?</p>	<p>11.</p> <p>a. What model of car is most popular for parents to drive in your city? </p> <p>b. What model of car do your parents drive?</p>

In order to answer statistical questions you must collect data. There are two types of data, *numerical* and *categorical*. Suppose you were interested in the Star Wars movies, and collected data on how much money each of the eight (so far) Star Wars movies made. This would be numerical data because each of the eight data values would be a numerical value. If you collected data on who your favorite character was in each movie this would be categorical data. Each of the seven data values would not be a number but a name of a character, which you could think of as a category.

**There are two types of data, *numerical* (sometimes called *quantitative*) and *categorical* (sometimes called *qualitative*). In a numerical data set every value is a number that represents a count or measurement. In a categorical data set every value is non-numerical.**

12. Identify each of the following data sets as categorical (C) or numerical (N).

- a. How far can a 6<sup>th</sup> grader jump?
- b. How many minutes a day does each person in our class read?
- c. What did each person in our class eat for lunch?
- d. How many pencils are in each person's desk?
- e. What is the favorite football team of each of the teachers at our school?
- f. How many letters are in each of our last names?
- g. Can you roll your tongue?
- h. In what zip code do students in our class live?

13. For each of the statistical questions below, determine if the data is categorical (C) or numerical (N). List three possible data values for each question.

- a. How much do boxes of cereals cost at the grocery store?
- b. What is the most popular kind of cereal sold at the grocery store?
- c. At which grocery store do people in your neighborhood like to buy cereal?
- d. How many boxes of Cheerios were sold at the store each day in the month of January?

**Directions:** Discuss why each question given below is, or is not, a statistical question. Write “yes” if it is statistical or “no” if not. For the questions that are not statistical, change them so that they can be statistical questions. Then determine if the data collected to answer each question is categorical (C) or numerical (N).

14. How long does it take you to travel to school every day?	15. How many boxes of cookies did each Girl Scout sell?
16. What is your resting heart rate at this moment?	17. How many jumping jacks can a kindergarten student do in 30 seconds?
18. Which brand of bubble gum holds its flavor the longest?	19. What is the typical number of cars that each family owns on your street?
20. How many text messages did you send yesterday?	21. Which video game do 6 <sup>th</sup> graders like to play?
22. How many cars in your town were sold this month?	23. How many days have people in my class been absent from school so far?
24. How many cars were sold each day this month?	25. How many pets does each of the students in my class have?
26. How old is Caden’s pet?	27. What did you eat for lunch yesterday?



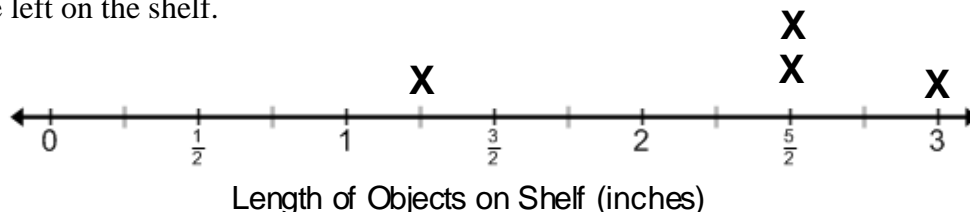
28. What pizza toppings do the people in my class like?	29. What is your teacher's favorite type of pizza?
30. How many pockets do I have on the clothes I am wearing today?	31. What is the speed of cars driving in front of our school?

32. Your class is having an ice cream party. Create two statistical questions that could help make your ice cream party a success. Then determine if the data collected to answer each question is categorical or numerical data.

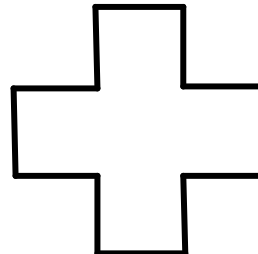
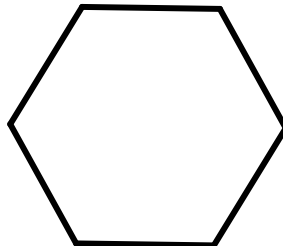
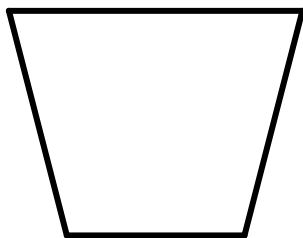


### Spiral Review

1. Barbara is measuring several objects to line up on a small shelf in her room. The shelf is 10 inches long. She has recorded the length of each object on the line plot below. Use the information given in the line plot to determine if Barbara's shelf is long enough to fit all of the objects, and if so, how much room will she have left on the shelf.



2. Find and draw all lines of symmetry for the following figures. Write the number of lines of symmetry that you found next to each figure.





#### 4.1a Homework: What is a Statistical Question?

Yesterday, Ruth and Carl invited 10 friends to go out to lunch. The questions below came up during the meal. Decide whether or not each question is a statistical question, and justify your decision.

1. How much does each person's meal cost?
2. How long did it take for each person to get to the restaurant?
3. Would Carl rather have burgers or pizza?
4. What percent of the bill did each person leave for a tip?
5. What was the most frequently ordered dish in the restaurant during lunch time?
6. Add one more statistical question of your own that relates to this situation. Justify why it is a statistical question.

Directions: Given each pair of questions below, circle the question that is a statistical question and then explain why.

<p>7.</p> <p>a. On a scale from 1 to 10 how much do you like math?</p> <p>b. On a scale from 1 to 10 how much do the people in your class like math?</p> 	<p>8.</p> <p>a. How long will it take to drive to the campground?</p> <p>b. How long did it take you to drive to the campground on Saturday?</p> 
--	--

9. Identify each of the following data sets as categorical (C) or numerical (N).

- How old are the grandmas for people in our class?
- On which day of the month were you born?
- How many pieces of paper are in each person's backpack in our class?
- What color of shoes are people in my class wearing today?
- What is the favorite subject in school for people in your class?
- How many marshmallows can you eat in 30 seconds?
- Can you do a flip on a trampoline?
- What is the area code for people's cell phone numbers in your class? **C**

10. Create a statistical question that could be answered by collecting numerical data.

11. Create a statistical question that could be answered by collecting categorical data.

**Directions:** Discuss why each question given below is, or is not, a statistical question. Write “yes” if it is statistical or “no” if not. For the questions that are not statistical, change them so that they can be statistical questions. Then determine if the data collected to answer each question is categorical (C) or numerical (N).

12. How much does a gallon of gas cost?	13. Who is Mr. Reed’s favorite NFL football player?
14. How many laps can a 6 <sup>th</sup> grader run during P.E. time?	15. How many songs do you have on your MP3 player?
16. What is the typical number of TVs that each person in your class has at home?	17. Which brand of AAA batteries lasts the longest?
18. How much does your backpack weigh?	19. How many cans of orange soda did the grocery store sell each day this week?

20. You are starting a small business mowing lawns in your neighborhood for the summer. Create two statistical questions that will help make your lawn mowing business a success. Then determine if the data collected to answer each question is categorical or numerical.



#### 4.1b Class Activity: Creating and Analyzing Dot Plots



Each member of your class is going to make and fly a paper airplane using the instructions given below or your own method. Write one statistical question that you could ask in relation to the paper airplanes. Be ready to share your question with the class.

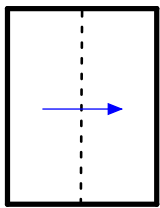
1. My statistical question about the paper airplane is, \_\_\_\_\_

---

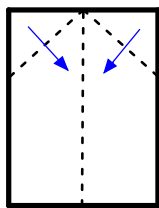


#### Instructions for folding a Paper Airplane

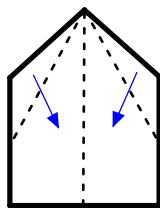
Make an airplane from a piece of paper that measures  $8\frac{1}{2}$  by 11 inches.



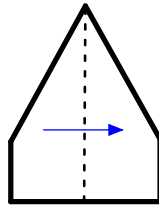
1. Fold paper in half and then unfold.



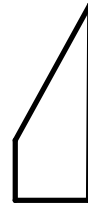
2. Fold the top two corners.



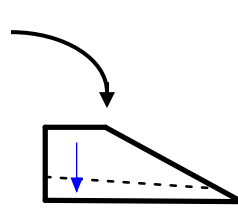
3. Fold corners again.



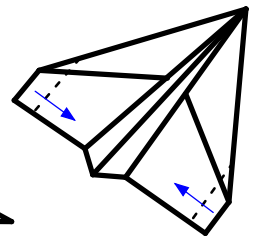
4. Fold paper in half again.



5. Rotate your paper 90 degrees.



6. Fold wings out on both sides.



7. Fold wing tips up.

We are going to try and answer the statistical question of,

**“How far will a paper airplane fly?”**

Predict how far you think the typical paper airplane made in your class will fly.

2. Prediction: \_\_\_\_\_

3. What units of measurement do you think the class should use and why?

Follow the instructions from your teacher for gathering the data.

4. Copy the dot plot made by the class below by using a dot to represent each data value.

A **dot plot** uses a number line to show the number of times each value in a data set occurs. A dot plot shows the ***distribution*** of the data set. Distribution describes how the data is related and organized; it can be described by its ***center***, ***spread (variability)***, and overall ***shape***. We are going to discuss how to interpret the shape of the data set using words like **peak**, **gap**, **cluster**, **symmetrical**, **skewed**, and **outlier**.



5. How close was your prediction?

6. The most recent test scores for Mr. Petrov's science class are shown in the table. Make a dot plot of the data. Be sure to label your number line and give it a title.

Mr. Petrov's Test Scores					
79	91	72	89	84	77
71	82	80	76	91	83
80	79	82	81	78	78
82	72	86	81	80	64
83	79	82	84	81	81
92	80	80	87	83	81

g. Make a dot plot of the data. Be sure to label your number line and give it a title.

h. How many students are in Mr. Petrov's science class?

i. Which test score was earned by the largest number of students?

j. Describe any peaks, clusters, or gaps in the data by marking them on the plot.

k. What is the overall shape of the data? Justify your answer.

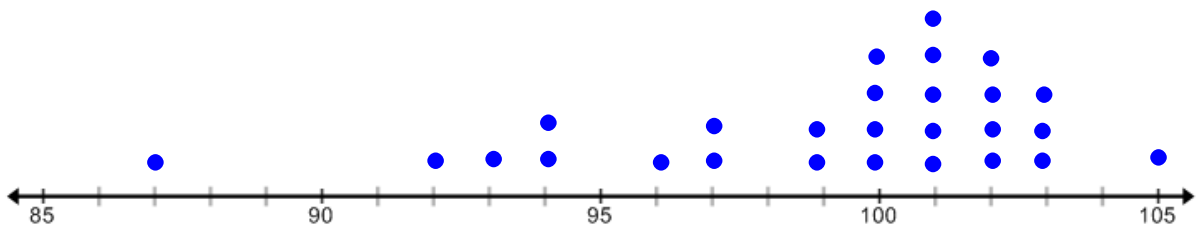
l. Mr. Petrov asks, "What is the most typical score for this test?" Use the distribution of data to answer this question.







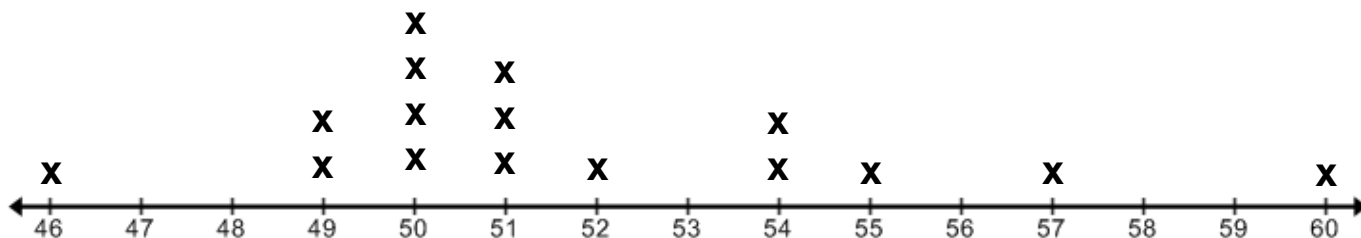
7. Marta records the high temperatures for each day she goes swimming in the month of August. She has recorded her data on the dot plot below.



- Marta did not label her dot plot with units or a title. Determine the appropriate units for this data and how the data was collected. Then give the dot plot an appropriate title.
- How many days did Marta attend the pool in August?
- Mark any clusters, gaps, or peaks that you observe on the graph.
- What is the overall shape of the data?
- Write and answer a statistical question related to the data shown in the plot.

8. The numbers below represent the amount of time in hours that several 6<sup>th</sup> grade students spent doing homework last week.
- 5, 15, 10, 11, 11, 6, 8, 0, 6, 13, 11, 7, 7, 12, 7, 12
- a. What attribute is being measured in the data above?
  - b. What is the unit of measurement for the attribute?
  - c. How could this data have been collected?
  - d. From how many people do we have data for?
  - e. Create a statistical question that could be answered using this data.
  - f. Plot the data on a dot plot, be sure to label the units and give it an appropriate title.
  - g. Mark any clusters, gaps, peaks or outliers that you observe on the plot.
  - h. What is the overall shape of the data?
  - i. Use the dot plot to answer the question “What is the typical amount of time that a 6<sup>th</sup> grader spends on homework each week?”
  - j. If you increased the number of students surveyed how would the graph change?
  - k. This data represents 6<sup>th</sup> graders. Do you think the answer to our question would change if we included high school students? Why or why not?

9. Which statistical question best matches the line plot given below? (Note: A line plot is just like a dot plot but instead of using dots for each data entry it has an **x**.) Once you have chosen a question, determine an appropriate unit to measure the data related to this question.



- How many glasses of water do each member of our class drink a day?
- How many minutes does it take to eat a Popsicle?
- How fast do cars travel down the highway in our town? Units: Mile per hour.
- How many gallons of water does a 6<sup>th</sup> grader use when showering?

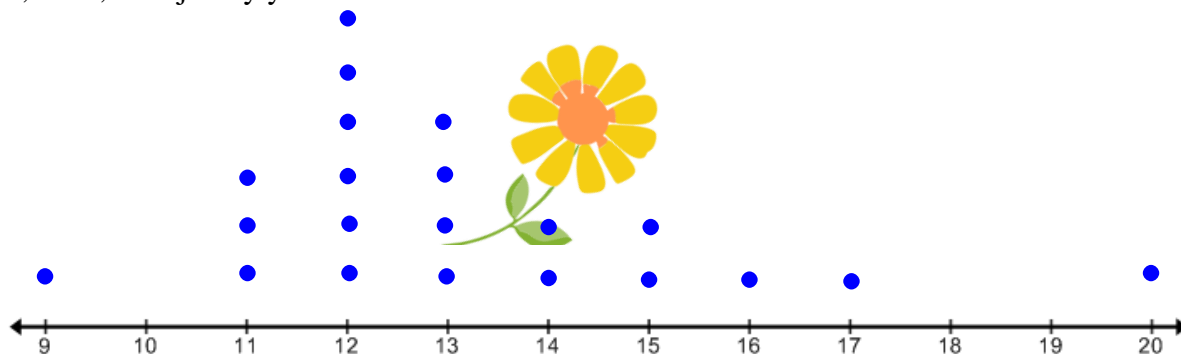
10. Write a title for this plot based off of the statistical question that you chose in number 9.

11. Use the shape of the plot to answer your chosen statistical question.

12. Find, Fix, and Justify



The dot plot below is about the number of yellow flowers that are in each vase at a wedding. Penelope is interpreting the data and makes the following statements. For each statement given below, find her mistake, fix it, then justify your answer.



Number of Yellow Flowers in Vases

- “The dot plot shows that 13 vases have 4 yellow flowers each.”

b. “The data appears to be skewed left.”






c. “The dots above the number 12 are outliers because there are more of them compared to the other numbers.”

## Spiral Review

The tally chart shows a survey of students’ favorite sports.

Sport	Tally
Soccer	
Football	
Basketball	
Volleyball	
Swimming	

1. Use the chart to complete the picture graph below. The first one has been done for you.

Favorite Sports				
				
Soccer	Football	Basketball	Volleyball	Swimming

2. How many students were surveyed?

3. The same number of students picked \_\_\_\_\_ and \_\_\_\_\_ as their favorite sport.

4. How many students picked volleyball as their favorite sport?

5. How many more students chose soccer than swimming as their favorite sport?

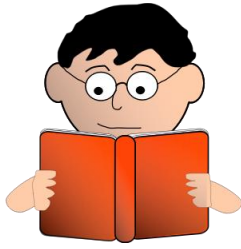
4.1b Homework: Creating and Analyzing Dot Plots

- 1. Create a statistical question of your own that you can answer by surveying people in your class. Then ask at least 15 people in your class your question and record your data below.
  - a. Write your question here:\_\_\_\_\_
  - b. Record your data in the table below.

Student	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
Response															

- c. What is the attribute that is being measured from your question?
  - d. What is the unit of measurement for your attribute?
  - e. Display your data in a dot plot. Be sure to label your number line and give it a title.
  - f. Mark any clusters, gaps, peaks or outliers that your observe on the plot.
  - g. What is the overall shape of the data?
  - h. Use the distribution of the data in your dot plot to answer your question.

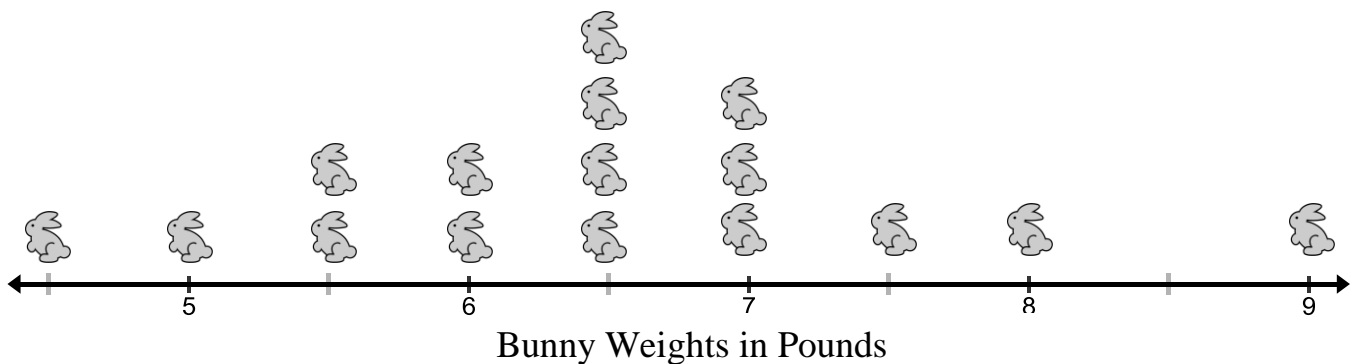
2. The table given shows how much time it took for 3<sup>rd</sup> graders to read the same book.



Student	Time (minutes)	Student	Time (minutes)
1.	4	6.	6
2.	7	7.	8
3.	8	8.	5
4.	14	9.	8
5.	6	10.	9

- What attribute is being measured in the table.
- What is the unit of measurement for this attribute?
- How was this data measured?
- Display the data in a dot plot, be sure to label the number line and give it a title.
- Mark any clusters, gaps, peaks or outliers that you observe on the plot.
- What is the overall shape of the data?
- What is the most typical time that it takes for a 3<sup>rd</sup> grader to read this book?


3. Use the graph below to answer the following questions.



- How many observations are there?
- What attribute is being measured in the dot plot?

- c. What is the unit of measurement for this attribute?
- d. How was the attribute measured?
- e. Mark any clusters, gaps, peaks or outliers that you observe on the plot.
- f. What is the overall shape of the data?
- g. What is the typical weight for this kind of bunny?

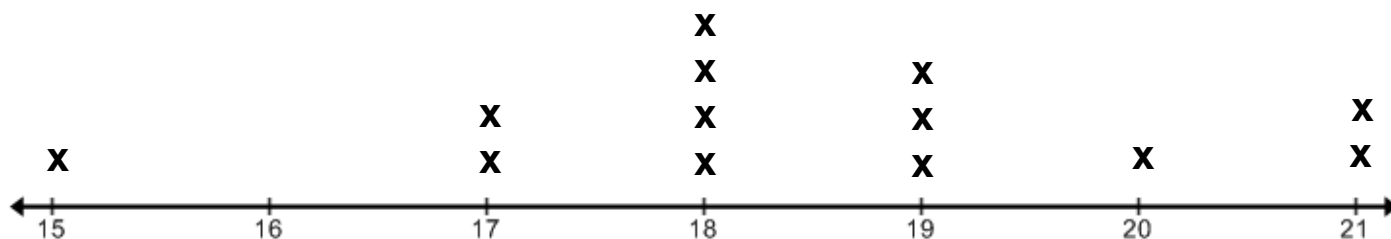
4. Use the data given below to answer the questions that follow.

100 Yard Dash Times (seconds)		
		
16.3	16.7	16.4
16.4	16.9	16.4
17.1	17.4	16.6
16.5	16.9	17.0
16.5	16.6	16.5

- a. Display the data above in a dot plot.
- b. How many people ran in the race?
- c. Circle the statement below that **does not** accurately reflect the data above.
  - The data in the plot is skewed right.
  - The attribute being measured is distance.
  - The typical time for runners in this race was around 16.5 seconds.
  - There is a cluster around the values of 16.9 seconds through 17.1 seconds.



5. Which statistical question best matches the line plot given below? (Note: A line plot is just like a dot plot but instead of using dots for each data entry it has an x.) Once you have chosen a question determine an appropriate unit to measure the data related to this question.



- a) How many siblings do students in my class have?
  - b) What is the typical height of flowers in my garden?; Units:Inches
  - c) How many hours of sleep do students in my class get each night?
  - d) How much money do people in my city spend on rent each month?
6. Write a title for this plot based off of the statistical question that you chose in number 5.
7. Use the shape of the plot to answer your chosen statistical question.

### 4.1c Class Activity: Creating and Analyzing Histograms

The high scores for a popular video game are shown below. Li wants to know what the typical high score for this video game is.

81	81	90	78	62	84	92	66	77	6	65	53	51	60
19	10	79	71	71	59	51	69	21	35	44	52	50	58
39	33	27	38	35	50	105	71	60	48	57	42	47	45
27	66	22	68	56	45	41	30	64	63	65	60	59	59

Li begins to make a dot plot for this data but soon realizes that a dot plot is not the best graphical display for this data. Discuss why a dot plot may not best for this data and see if you can come up with a better way to represent the data.



1. Copy the frequency table you made as a class below.

A **frequency table**, or tally chart, groups data into intervals. The *frequency* is the number of values in an interval.

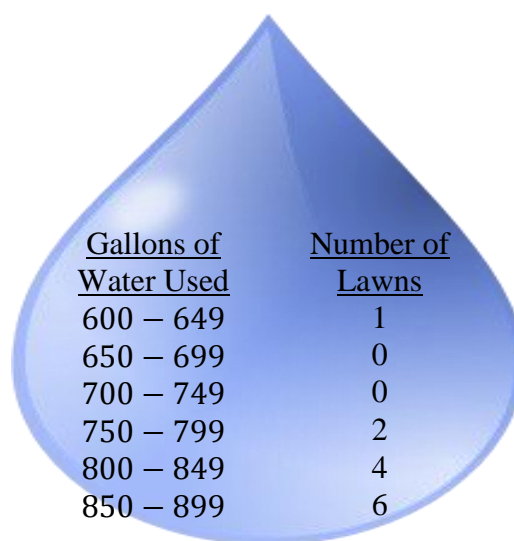
2. Copy the histogram you made as a class. Be sure to note key features on the graph.

A **histogram** is a frequency display that uses bars to show the distribution of data in a set. The height of the bar shows the frequency of the values in that interval.



3. The graphic below shows amount of water used on a particular day for people's lawns in the same neighborhood.

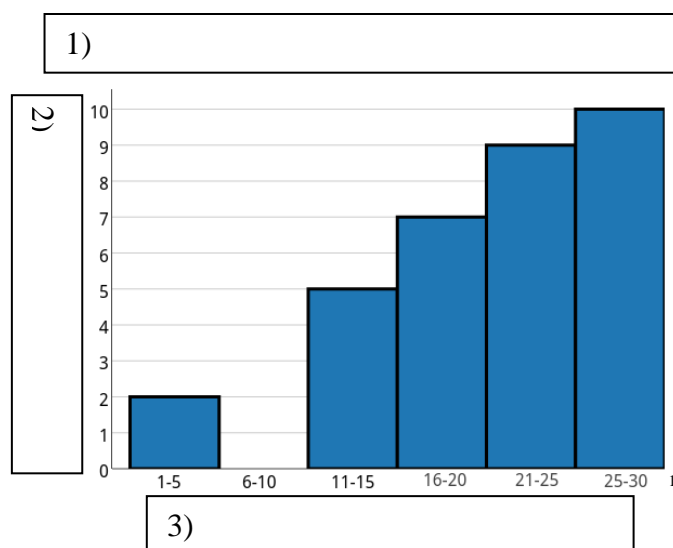
- What is the attribute being measured?
- What is the unit of measurement for this attribute?
- How would you measure this data?
- Display the data in a histogram?



<u>Gallons of Water Used</u>	<u>Number of Lawns</u>
600 – 649	1
650 – 699	0
700 – 749	0
750 – 799	2
800 – 849	4
850 – 899	6

- Mark any clusters, gaps, peaks or outliers that you observe on the plot.
- What is the overall shape of the data?
- What is the typical amount of water used on a lawn in this neighborhood?

4. The graph below shows the number of Apps that my friends have on their electronic tablets.



- Write in appropriate labels for boxes 1 through 3.
- How many of my friends have between 21 and 25 apps on their tablets?
- How many of my friends have between 6 and 10 apps on their tablets?
- What does the shape of the distribution tell you?
- How many of my friends were surveyed?
- Mandy has 17 apps on her tablet. She asks where her particular data value is on the graph. Is it possible to show her this?

5. Marissa is interested in the number of days that students in her class are absent each year. On the last day of school, she asks several students in her class how many times they missed school for the entire year. Her data is given below.

6	12	0	2	16	5	4	0	5	5
1	7	4	21	3	13	2	3	9	5

- a. Make a histogram of this data. (If needed, make a frequency table first.)

- b. Circle the statement below that cannot be made about the data displayed in your histogram?

One student was absent at least 20 days and at most 24 days.

The data in the histogram is skewed left

Marissa surveyed 20 students.

Less than  $\frac{1}{4}$  of the class was absent more than 10 days.

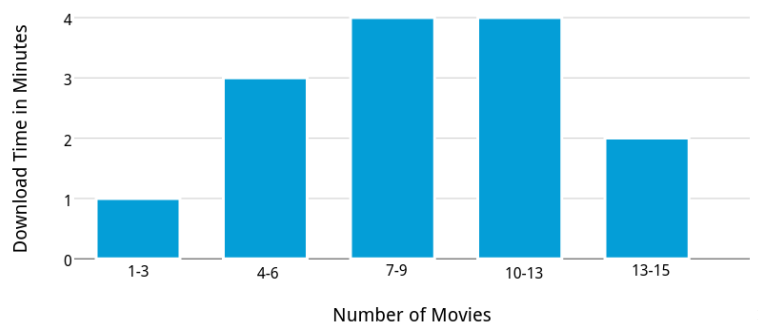
6. Find, Fix, and Justify.

The table given below shows the times that it takes to download different movies.

Time to Download a Movie in Minutes	1 – 3	4 – 6	7 – 9	10 – 12	13 – 15
Number of Movies	1	3	4	4	2

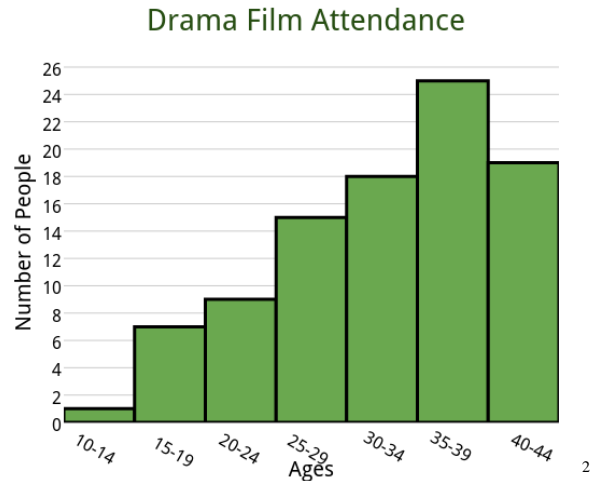
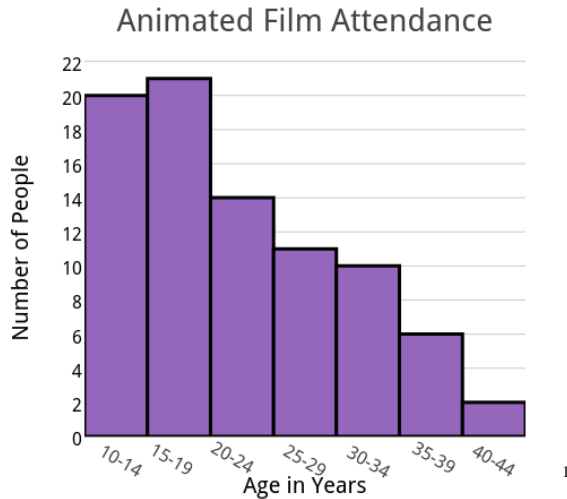
- a. Carmen made a histogram of the data and has made many mistakes. Find at least two of her mistakes, fix them on the histogram and justify why your reasoning is correct.

Movie Download Times



- b. Carmen has just purchased a movie online based on the data in the graph, what would you expect the typical download time to be for this movie?

7. The two histograms given below show the ages of people watching two different types of movies at the same multi-screen movie theater.



- Describe the shape of the distribution for each type of movie.
- How many people attended the drama?
- What percentage of the people attending the drama film were younger than 20?
- What percentage of people attending the animated film were 30 or older?
- Suppose you work for a marketing agency that is going to advertise these movies. Which type of movie would you advertise to a younger audience? Justify your answer.

8. The data in the table below shows the ring finger circumference of 10 people.

Ring Finger Circumference in Millimeters									
15.4	17.6	16.8	16.1	17.8	15.7	15.3	15.7	16.9	15.6

- a. Display the data in a histogram.

1

- a. Which interval includes the most data values?
- b. How many people have a ring finger circumference less than 16.5 mm?
- c. What percentage of people have a ring finger circumference greater than 17 mm and less than 18 mm?
- d. Describe the shape of the distribution.

### Spiral Review

1. Use long division to find each solution. Round your answer to the nearest tenth.
- a.  $25 \div 4 =$                                       b.  $515 \div 6 =$                                       c.  $732 \div 20 =$
2. How would your share...
- a. 12 pieces of licorice evenly between 5 friends?
- b. 58 jelly beans evenly between 7 friends?



4.1c Homework: Creating and Analyzing Histograms

1. Create a statistical question of your own that you can answer by surveying people in your class. Then ask at least 15 people in your class your question and record your data below.

a. Write your question here:\_\_\_\_\_

b. Record your data in the table below.

Student	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
Response															

c. What is the attribute that is being measured from your question?

d. What is the unit of measurement for your attribute?

e. Display your data in a histogram, be sure to label your number line and give it a title.

f. Mark any clusters, gaps, peaks or outliers that you observe on the histogram.

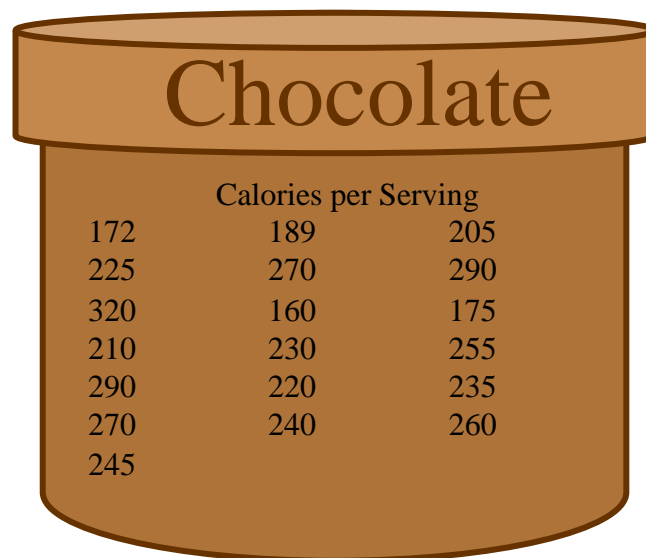
g. What is the overall shape of the data?

h. Use the distribution of the data in your histogram to answer your question.

2. A nutritionist is interested in the typical number of calories in a serving of chocolate ice cream. He has collected data for calories per serving for several different brands of ice cream. His data is shown below.

a. What attribute is the nutritionist measuring?

b. Make a frequency table of the data.



c. Make a histogram of the data.

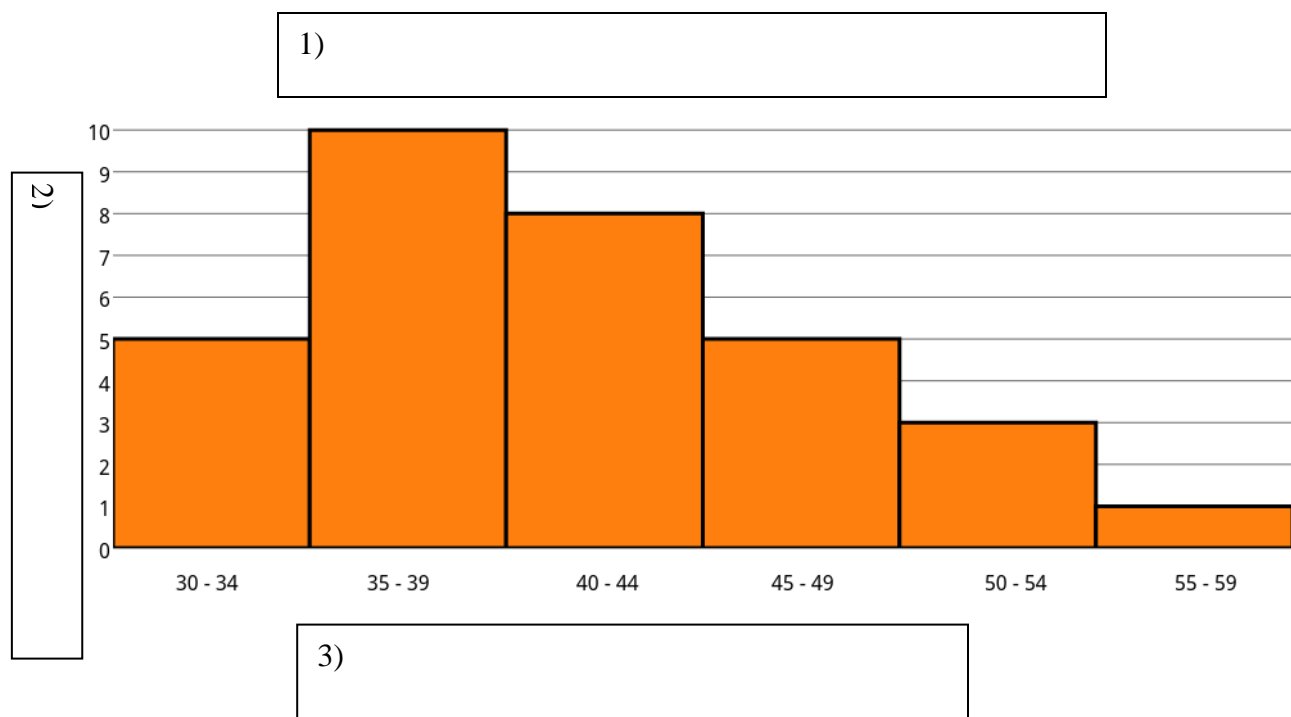
1

d. Mark any clusters, gaps, peaks or outliers that you observe on the histogram.

e. What is the overall shape of the data?

f. What is the typical number of calories in a serving of chocolate ice cream?

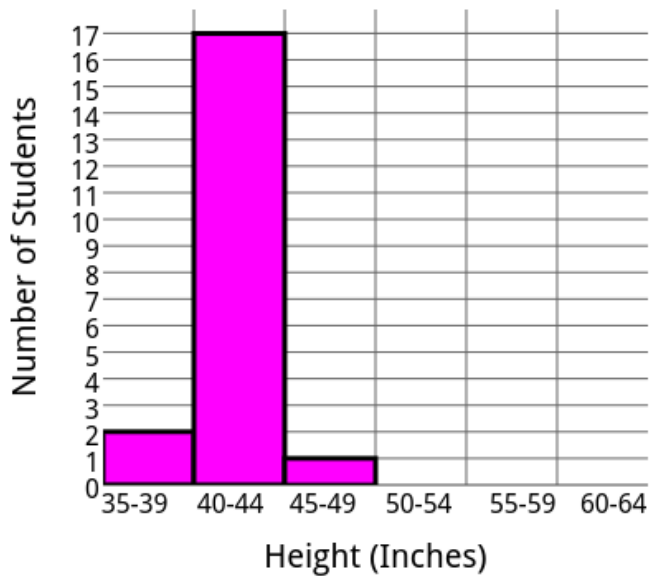
3. The histogram below shows the ages of parents for students in Carlotta's 6<sup>th</sup> grade class.



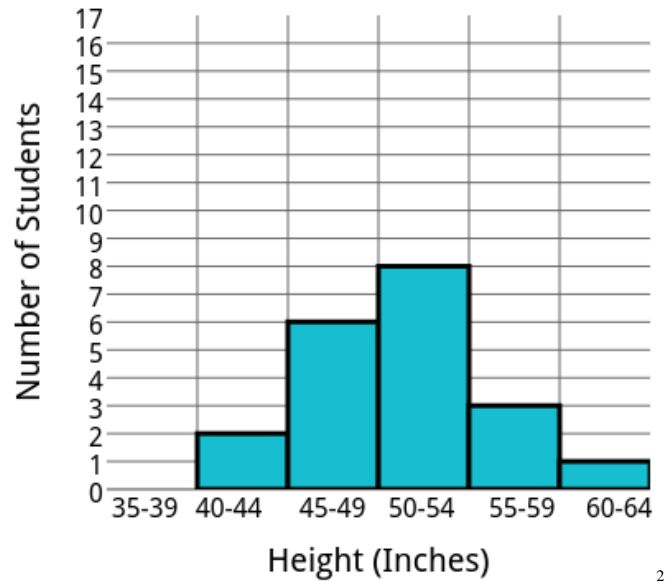
- Write in appropriate labels for boxes 1 through 3.
- How many students have a parent between the ages of 45 and 49?
- How many of students have a parent older than 49?
- What does the shape of the distribution tell you?
- How many students were surveyed?
- Carlotta wants to know how many parents are younger than 37, is it possible to determine this from the histogram? Explain.
- Create your own question that could be answered using the histogram. Then answer your question.

4. Ricardo and Sven are interested in the heights of students at their school. At the end of the day Ricardo stands at the front door of the school and collects the heights of 20 students leaving the school. Sven collects the heights of 20 students leaving the kindergarten classroom. They each make a histogram of the data they collected.

Title:



Title:



a. Determine which histogram belongs to each boy by giving each histogram a title.

b. What attribute is being measured in the histograms.

c. What unit of measurement for this attribute?

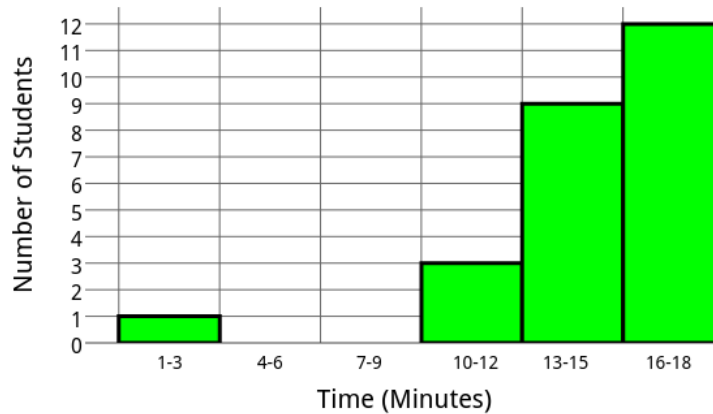
d. How was this data measured?

e. What percent of kindergartener's are less than 40 inches tall?

f. What percent of students are at least 50 inches tall?

5. Use the histogram below to answer the questions that follow.

Time for 3rd Graders to Complete a Puzzle



- a. What attribute is being in the histogram, what is the unit of measurement for this attribute, and how was this data measured?
- b. Circle all the statements below that are true about the distribution of the data in the histogram.
  - i. 25 third graders were surveyed.
  - ii. The data is skewed left.
  - iii. It took between 16 and 18 students 12 minutes to complete the puzzle.
  - iv. It took one student between one and three minutes to complete the puzzle. This student is an outlier.
  - v. It takes the typical 3<sup>rd</sup> graders between 13 and 18 minutes to complete the puzzle.
- c. Suppose you gave a class of 6<sup>th</sup> graders the same puzzle and recorded how much time it took each person to complete the puzzle. Describe how the 6th grade histogram would be different than the 3<sup>rd</sup> grade histogram.

#### 4.1d Self Assessment: Section 3.1

Consider the following skills/concepts. Rate your comfort level with each skill/concept by checking the box that best describes your progress in mastering each skill/concept. Corresponding sample problems, referenced in brackets, can be found on the following page.

<b>Skill/Concept</b>	<b>Minimal Understanding 1</b>	<b>Partial Understanding 2</b>	<b>Sufficient Mastery 3</b>	<b>Substantial Mastery 4</b>
1. Recognize a statistical question as one that anticipates variability in the data.	I can recognize a statistical question in some contexts.	I can recognize a statistical question in a variety of contexts.	I can recognize a statistical question in a variety of contexts and create statistical questions on my own.	I can recognize a statistical question in a variety of contexts and create statistical questions on my own. I can rephrase questions that are not statistical into statistical questions.
2. Identify if the set of data used to answer a statistical question is numerical or categorical.	I can identify if a statistical question will generate numerical data. I struggle to identify questions that will generate categorical data.	I can identify if a statistical question will generate numerical or categorical data.	I can identify if a statistical question will generate numerical or categorical data. I can suggest possible data values for a statistical question.	I can identify if a statistical question will generate numerical or categorical data. I can suggest possible data values for a statistical question. I can create questions that generate categorical and statistical questions.
3. Display numerical data in a dot plot, use key terms to describe its shape, and interpret the distribution of data.	I can display data in a dot plot but often don't know how to label it appropriately. I don't know how to describe the shape or interpret the distribution of data.	I can display data in a dot plot that is accurately labeled. I can identify some of the terms used to describe its shape. I struggle to interpret the distribution of data.	I can display data in a dot plot that is accurately labeled. I can identify and use terms such as peaks, clusters, gaps, skewed, symmetrical and outliers to describe the shape. If the data is skewed I cannot define its direction.	I can display data in a dot plot that is accurately labeled. I can identify and use terms such as peaks, clusters, gaps, skewed, symmetrical and outliers to describe the shape and interpret the distribution of data.
4. Display numerical data in a histogram, use key terms to describe its shape, and interpret the distribution of data.	I can display data in a histogram but often don't know how to label it appropriately. I don't know how to describe the shape or interpret the distribution of data.	I can display data in a histogram that is accurately labeled. I can identify some of the terms used to describe its shape. I struggle to interpret the distribution of data.	I can display data in a histogram that is accurately labeled. I can identify and use terms such as peaks, clusters, gaps, skewed, symmetrical and outliers to describe the shape. If the data is skewed I cannot define its direction.	I can display data in a histogram that is accurately labeled. I can identify and use terms such as peaks, clusters, gaps, skewed, symmetrical and outliers to describe the shape and interpret the distribution of data.

Skill/Concept	Minimal Understanding 1	Partial Understanding 2	Sufficient Mastery 3	Substantial Mastery 4
5. Summarize a numerical data set in relation to its context by reporting on the number of observations, identifying how the data was measured and its units of measurement.	When summarizing a data distribution I can identify the number of observations.	When summarizing a data distribution I can identify the number of observations and offer ideas on how the data was measured	When summarizing a data distribution I can identify the number of observations, offer ideas on how the data was measured, and identify appropriate units of measurement.	When summarizing a data distribution I can identify the number of observations, offer ideas on how the data was measured, and identify appropriate units of measurement. I look back and account for the context of a data distribution.

### Sample Problems for Section 4.1

*Square brackets indicate which skill/concept the problem (or parts of the problem) aligns to.*

1. Evan and Somer are selling cookies to the kids at their school during lunchtime for their school band fundraiser. They have a variety of cookies that each cost a different amount. The questions below were discussed as they worked on the fundraiser. [1] [2]

For each question decide if it is a statistical question and justify your decision. Then state whether the data collected to answer the question will be categorical or numerical data.

- a. What is the price of cookies for the fundraiser?
- b. How long did it take us to sell out of cookies on Tuesday?
- c. How many cookies were sold each day of the fundraiser?
- d. What is Evan's favorite cookie?

For each of the questions above that are not statistical, rephrase them so that they are a statistical question. Write them below. List three possible data values for each question.

e.

f.

Create two statistical questions of your own, one categorical and one numerical, which are related to the cookie fundraiser.

g. Numerical:

h. Categorical:

2. Beckham is interested in the hand span, the length from the tip of the pinky finger to the tip of the thumb, for people in his class. He asks the statistical question, “How long is the hand span for people in my class?” He decides to conduct a survey to answer his question. His data is shown below. [3][5]

Length of hand span in inches: 6, 6.5, 9, 7.5, 7.5, 7, 6.5, 7, 7, 9.5, 6, 8, 5.5, 7, 6.5

- a. What is the attribute that is being measured?
- b. What is the unit of measurement for this data?
- c. What tools do you think Beckham used to obtain these measurements?
- d. For how many people did Beckham measure the hand span?
- e. Display the data in a dot plot.
- f. What is the shape of the data? Be sure to use key terms in your description.
- g. Use the shape of the data to interpret the distribution of data and answer Beckham’s question.



3. The time that it takes for several 6<sup>th</sup> graders to eat their lunch is shown below. [4][5]

How long does it take you to eat your lunch? (minutes)					
26	24	21	25	19	26
26	24	16	13	15	21
27	27	14	27	22	20
24	4	25	29	27	22
20	17	28	22	26	25

- What statistical question could be asked that relates to the data.
- What is the attribute that is being measured?
- What is the unit of measurement for this data?
- What tools do you think Beckham used to obtain these measurements?
- How many people were surveyed?
- Display the data in a histogram.

1

- How many people spent 16-20 minute eating lunch?

- h. What percentage of people spent 15 minutes or less eating lunch?
  
  
  
  
  
  
  
  
  
  
- i. What is the shape of the data? Be sure to use key terms in your description.
  
  
  
  
  
  
  
  
  
  
- j. Use the shape of the data to interpret the distribution of data and answer your statistical question.

## Section 4.2: Measures of Center

### Section Overview:

In this section, the focus is on finding and interpreting measures of center. To be useful, measures of center must be characterized by numerical descriptions. The first lesson has students begin to understand the mean is a measure of center that levels out the data in the sense of a unit rate. In this “leveling out” sense the mean, (often referred to as the “arithmetic average” or simply “average”) is often thought of as a fair share. They abstract the act of gathering all the data and equally redistributing it to summing all the values in the data set and then dividing by the total number of items. In addition to looking at the mean as a fair share, students will experiment with the mean as a balancing point of the distribution. This prepares students to better understand mean absolute deviation when they analyze variability in the next section. Students practice finding the mean from several different data displays and begin to recognize that the mean can be affected by extreme values. Next, students learn about the median and mode and how to find them from a given set of data. They soon realize that they can choose a measure of center that best represents the data by analyzing the shape of the data distribution and the context in which the data is gathered.

### Concepts and Skills to Master:

*By the end of this section, students should be able to:*

1. Understand the mean as a “leveling out” of data or a fair share. Also understand how the mean can be interpreted as the “balancing point” for the data values.
2. Find the mean from a set of data and interpret its meaning for a given context.
3. Find the median or mode from a set of data and interpret their meaning for a given context.
4. Determine the best measure of center for a given data set and justify why it is the best measure of center.

#### 4.2a Class Activity: How Much is a Handful?

Directions: Grab a handful of linking cubes from your teacher. Count the number of cubes in your hand and write the number on a sticky note.

Recall that the distribution of a data set can be described by its *center*, *spread (variability)*, and *shape*. In this section we are going to investigate how to describe the **center** of a data set.

1. What do you think is the center of the class's linking cube data?
2. What if we wanted everyone in the class to have a fair share of cubes, how many cubes would there be per person?

Work in your group to answer the question above, be ready to discuss your ideas with the class.

This number that describes the data as a “fair share” or a “leveling out” of the data is called the arithmetic mean or just **mean**. It is often referred to as the **average**. It is a single number that describes the center of the data.

3. At the arcade it costs one token to play a game of Skee-ball. The table below shows the number of tokens that a group of 6 friends brings with them to the arcade.

Tokens					
Kira	Tom	Jade	Braden	Nan	Lonny
5	6	3	3	2	5

They decide that they want each person to have the same number of turns playing the game. Determine the number of times that each person can play the game. In other words, what is the **mean** or **average** number of times that each person can play the game? Use your linking cubes to represent the tokens if needed.

4. A family of five goes out to lunch. The numbers below show the cost of each item that a person ordered for lunch. They decide to split the bill so that everyone pays the same amount. In other words, what is the average cost per person for lunch?

\$8.25      \$8.70      \$7.50      \$6.95      \$9.35

Discuss how to find the mean for any given set of data.



**The *mean* or *average* is a *measure of center* for a data set. The mean is found by taking the sum of all the data values and then dividing by the number of data values.**

5. Find the average for each set of data.

a. Time to run a mile in minutes: 9.5, 8.7, 9.3, 10, 10.4, 7.5, 8.4, 9.1	b. Number of movies attended this summer: 4, 3, 9, 5, 4, 0, 7, 5, 6, 2, 3, 4																								
c. <table border="1"> <thead> <tr> <th colspan="2">Number Of Posts On A Social Media Site Per Week</th></tr> </thead> <tbody> <tr> <td>JJ 25</td><td>Kaci 48</td></tr> <tr> <td>Charlie 34</td><td>Mark 38</td></tr> <tr> <td>Andrew 56</td><td>Nicole 52</td></tr> <tr> <td>Quinn 34</td><td>Seth 46</td></tr> <tr> <td>Cindy 45</td><td>Roxy 30</td></tr> </tbody> </table>	Number Of Posts On A Social Media Site Per Week		JJ 25	Kaci 48	Charlie 34	Mark 38	Andrew 56	Nicole 52	Quinn 34	Seth 46	Cindy 45	Roxy 30	d. <table border="1"> <thead> <tr> <th colspan="2">Number of Pets Owned</th></tr> </thead> <tbody> <tr> <td>Zara</td><td>   </td></tr> <tr> <td>Alina</td><td> </td></tr> <tr> <td>Todd</td><td><del>    </del></td></tr> <tr> <td>John</td><td><del>    </del>  </td></tr> <tr> <td>Aarav</td><td>   </td></tr> </tbody> </table>	Number of Pets Owned		Zara		Alina		Todd	<del>    </del>	John	<del>    </del>	Aarav	
Number Of Posts On A Social Media Site Per Week																									
JJ 25	Kaci 48																								
Charlie 34	Mark 38																								
Andrew 56	Nicole 52																								
Quinn 34	Seth 46																								
Cindy 45	Roxy 30																								
Number of Pets Owned																									
Zara																									
Alina																									
Todd	<del>    </del>																								
John	<del>    </del>																								
Aarav																									

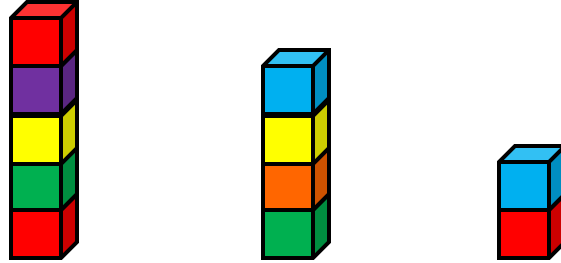
### Spiral Review

- Out of a class of 25 people, 14 of them own a dog. What percentage of the class owns a dog? Round your answer to the nearest percent.
- 6 out of 7 days a week Lola works out. What percentage of the week does Lola work out? Round your answer to the nearest percent.
- Carson got 16 out of 15 problems correct on a quiz. What percentage of the problems did he get correct? How could this be possible? Round your answer to the nearest percent.
- Three friends earned \$21.75 doing jobs for a neighbor. How much money did each person earn individually if they divide the total earnings equally?
- How many  $\frac{1}{3}$  cup servings are in 3 cups of nuts?

#### 4.2a Homework: Mean as a Fair Share

Find, Fix, and Justify

Tina and Arnold are trying to find the average number of blocks per stack for the linking cubes shown below. They have both made a mistake in their reasoning. Find their mistake and fix it, then justify why your thinking is correct.



1. Tina knows that she can think of an average or mean as every group having a fair share or the same number of blocks. She begins to rearrange the blocks and concludes that there are too many cubes for 3 groups of 3 and not enough cubes for 3 groups of 4, so there must not be a mean.
2. Arnold knows that he can find the average by adding up the total number of blocks and then divide by the number of groups. When he does this he gets an answer of  $3\frac{2}{3}$ . He claims that this does not seem like the right answer because the original set of numbers does not include  $3\frac{2}{3}$ .

## 4.2b Class Activity: Finding the Mean

Recall that the mean or the average is a measure of center that summarizes a data set.



**The *mean* or *average* is a *measure of center* for a data set; (a number that best describes the center of a data set). It found by taking the sum of the data values and then dividing by the number of data values.**

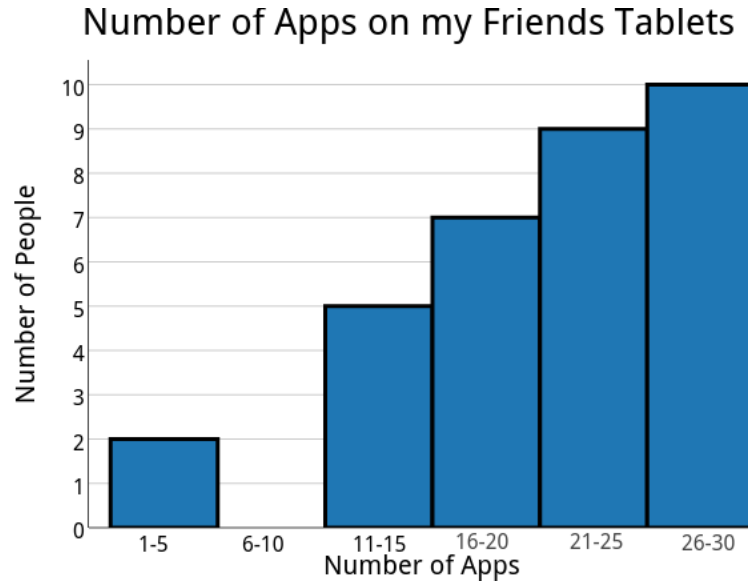
The most recent test scores for Mr. Petrov's science class are shown in the table.

Mr. Petrov's Test Scores					
79	91	72	89	84	77
71	82	80	76	91	83
80	79	82	81	78	78
82	72	86	81	80	64
83	79	82	84	81	81
92	80	80	87	83	81

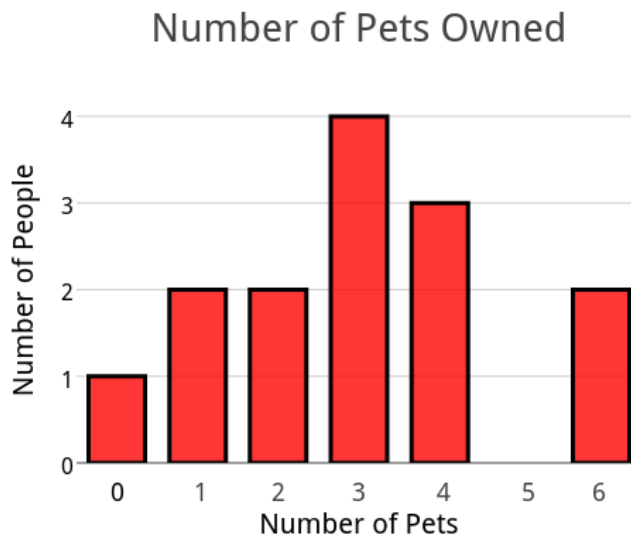
1. Find the mean test score for the class.
2. What does this tell you about how well the students in Mr. Petrov's class did on the test? Compare this answer with the conclusions you made about the dot plot that you made for this data in **Section 4.1b**.
3. Find the mean of your paper airplane data in the dot plot from **Section 4.1b**. Compare this average to your conclusions about how far a typical paper airplane will fly based on the distribution of the data in the dot plot.



4. The histogram below shows the number of apps that a large group of friends have on their electronic tablets. Is it possible to find the exact mean from the data presented in a histogram? Why or why not?

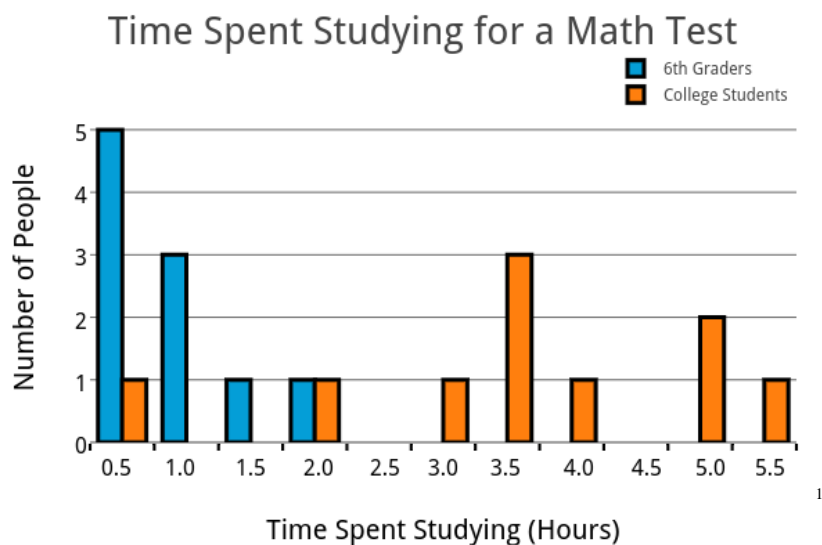


5. The **bar graph** below shows the number of pets that fifteen 6<sup>th</sup> graders have owned in their lifetime. Find the mean number of pets owned from the information given in the graph.

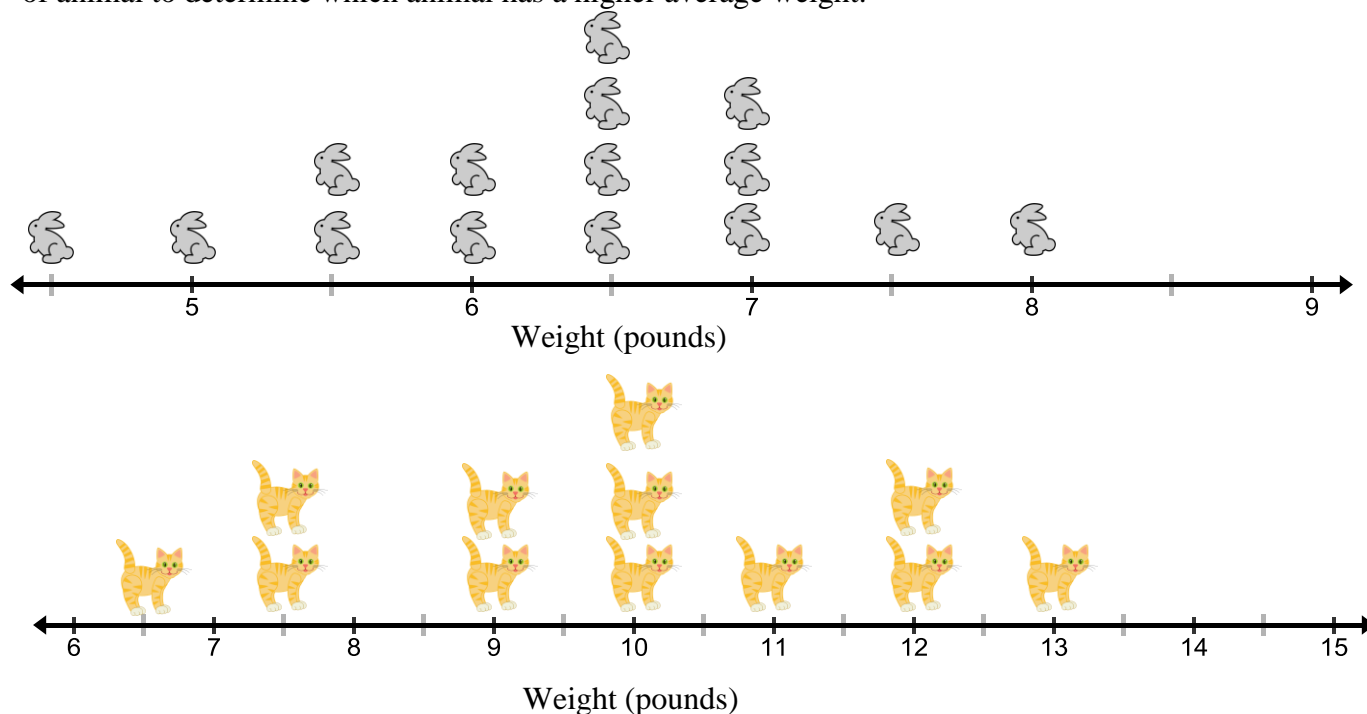


6. In your class 8 people do not get any weekly allowance, 3 people get \$5, 5 people get \$7, 7 people get \$10, 6 people get \$12, 1 person gets \$14 and 1 person gets \$15. Find the average amount of weekly allowance for the class.

7. The double bar graph shows the time spent doing homework each night for a class of 6<sup>th</sup> graders and a class of college students.



- a. Use the bar graph to estimate the average time that each group of students spent studying for the test.
- b. Calculate the average time each group spent studying for the test. Use the averages to determine which group of students spent more time studying for their math test.
8. The graphs given below show the weights of two different animals. Find the mean weight of each type of animal to determine which animal has a higher average weight.

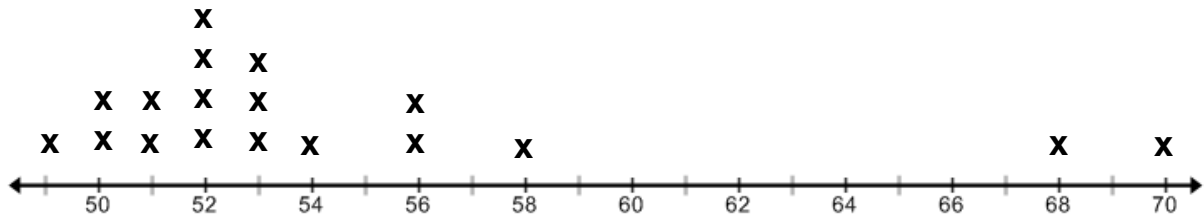


9. The table shows the low temperatures recorded for a week in early March.

Daily low temperatures in degrees Fahrenheit						
30°	33°	33°	30°	15°	32°	34°

- Identify any outliers in the data.
- Find the mean of the data with and without the outlier(s).
- Describe how the outlier affects the mean.

10. The line plot shows the speed of several cars traveling down the highway.



Speed of Cars on Highway in MPH

- Identify any outliers in the data.
- Find the average speed of the cars with and without the outlier(s).
- Describe how the outlier(s) affects the data.

11. Create two different sets of data that have 4 values each and a mean of 15.

12. Owen has recorded four of his quiz scores for his history class below.

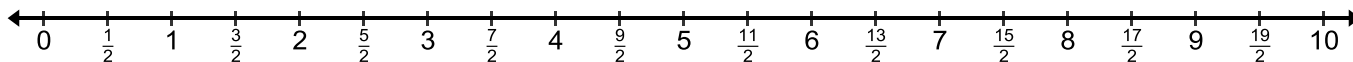
**7, 10, 8, 8**

- a. He has one more quiz to take and his mom has told him that if he has an overall average of at least 8 on all his quizzes he can earn extra time playing video games. Determine the lowest possible score he can get on his final quiz to earn his reward.
- b. Owen's history teacher wants to get an idea of how four of her students, including Owen, did on their 5 quizzes. The teacher's grade book is shown below. Assume that Owen gets a 7 on his last quiz. Show how Owen's teacher can find out how these 4 students did on their quizzes as a whole.

Student	Quiz #1	Quiz #2	Quiz #3	Quiz #4	Quiz #5
Leta	6	5	8	6	7
Owen	7	10	8	8	7
Seth	8	7	8	8	7
Somer	10	10	9	9	10

### Spiral Review

1. Use a ruler to find the length of four items in your desk to the nearest  $\frac{1}{4}$  of an inch. Then mark these measurements on the number line given below.



2. Solve the following problems without a calculator.
- a.  $34 \times 25$                       b.  $542 \times 33$                       c.  $432 \times 8332$

## 4.2b Homework: Finding the Mean

Directions: Find the mean for each data set. Round to the nearest tenth and be sure to include units in your answer if applicable.

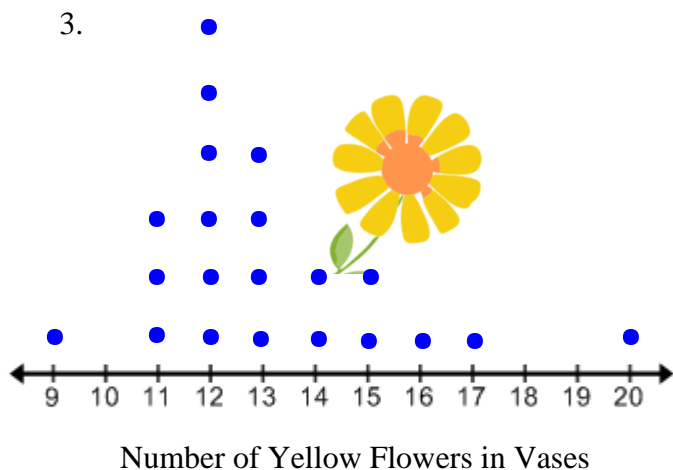
1.

Height of Players (cm)			
132	140	172	144
150	155	165	149
145	161	144	163

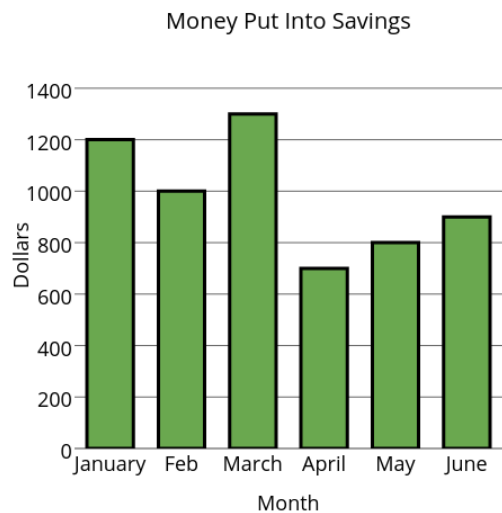
2.

Number of Boxes Ordered	
Lonna	
Gary	
Hunter	
Miriam	
Jonas	

3.



4.



5. Number of sit-ups: 30, 34, 20, 28, 31, 28, 27

6. Movies seen this summer: 4, 6, 3, 0, 2, 4, 1, 3, 2, 5, 4, 6, 3, 4, 2, 3, 1

7. 71, 67, 73, 73, 66

8. 4.5, 5.1, 4.6, 4.7, 4.5, 5.0, 4.9

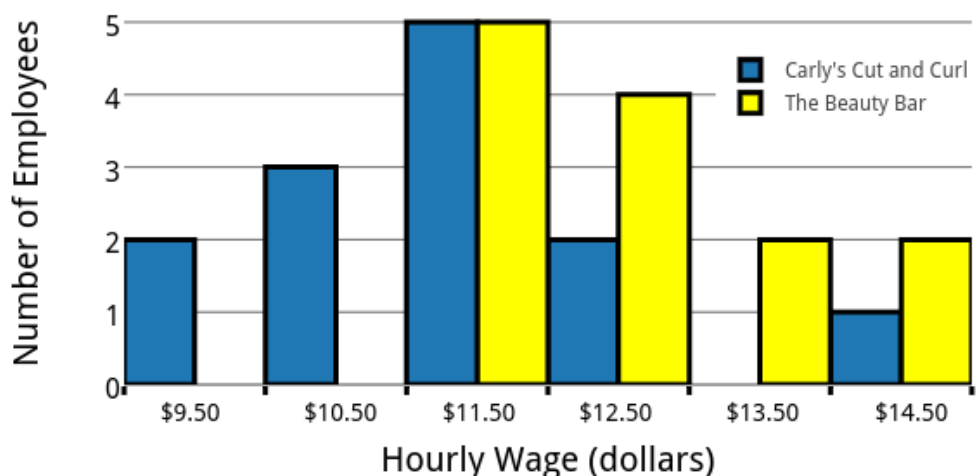
9. In order to qualify for the next round of a snowboarding competition, Tara must have an average score of at least 80 on her first 4 rounds of competition. Her scores for the first 3 rounds are below. What is the lowest possible score that she can get on her fourth round to move on in the competition?

Round 1	Round 2	Round 3	Round 4
90	75	72	



10. The bar graph below shows the hourly wage of employees that work at two different hair salons. Determine which salon will most likely pay you more by finding the average hourly wage for employees at each salon.

Hourly Wage of Employees at 2 Hair Salons



The data below shows the attendance of fans at each home game for the high school baseball team.

Fan attendance at home games: 34, 37, 30, 15, 34, 38, 41, 33

- Which data value is an outlier? Explain.
- Calculate the average attendance with and without the outlier. Then describe how the outlier affects the mean.  
Average Attendance with the outlier:  
  
Average Attendance without the outlier:
- Describe a situation that could have caused the outlier in this set of data.

12. The table below shows how much money Bethany spent on new school clothes.

Item	T-Shirt	Pants	Skirt	Shoes	Sweater	Shorts	Dress	Button – up Shirt
Cost	\$15.00	\$25.00	\$18.50	\$48.00	\$22.00	\$18.00	\$25.00	\$20.00

- a. Find the average cost per item?



- b. Do any outliers affect the average cost? If so, eliminate the outlier and find the new mean.
- c. Bethany insists that, on average, her friends spend more money per item on school clothes than she does. To be sure, she asks 6 of her friends to tell her the average amount of money per item that they spent on school clothes. Their data is given below.

Tami- \$13.50, Dusty- \$25.50, Wade- \$13.00, Bobby- \$28.30, Luis- \$17.00, Giselle- \$9.75

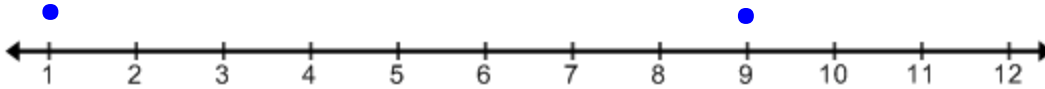
Find the average amount of money that Bethany's friends spend per item on school clothes and then compare it to her average cost per item.

#### 4.2c Class Activity: Mean as a Balance Point



Carly wants to know how long her friends can do a handstand. She asks two of her friends to do a handstand. One friend can do a handstand for 9 seconds, her other friend can only do a handstand for 1 second. She records their times on the dot plot below.

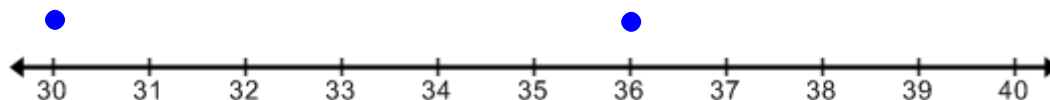
Number of Seconds My  
Friends Can Do A Handstand.



1. Carly's teacher says that the *center* of this data is the *balance point*. Discuss with a neighbor what you think she means by this. Then draw a triangle under the dot plot where you think the balance point is.
2. Carly notices that the balance point is also the mean of the two data values. Confirm this by calculating the mean below.
3. Can you place the values in a different location so that the plot still remains balanced at 5? Redraw the plot below with your new values.
4. Check that the plot is balanced by calculating the mean of your two new data values.



5. What do the **mean** and the **balancing point** of the data have in common?
6. Find the balancing point of the two data values on the plot below. Draw a triangle to indicate the point of balance.



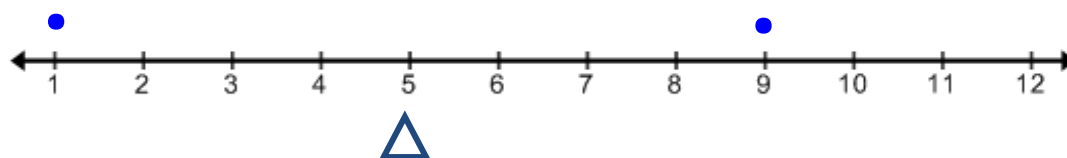
7. One data value of 7 is shown on the plot below. Where should the other value be placed so that the plot is balanced at 10?



8. If the dot at 7 is moved to the left two spaces, where should the other value be moved to so that the plot remains balanced at 10? Justify your answer.
9. Confirm that the mean of the new data points found in #8 is 10.

The original data for Carly's two friends have a balance point of 5. We justified that the balancing point was 5 because 1 and 9 are both 4 spaces away from 5. In mathematics we call these distances *absolute deviations* from the mean. The difference between the balancing point or mean and each data point is an *absolute deviation*. This is shown with Carly's handstand data below.

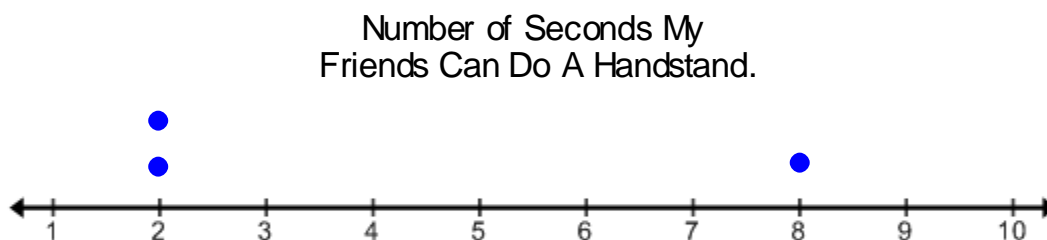
Number of Seconds My  
Friends Can Do A Handstand.



10. What relationship does the absolute deviation on the right of the balance point have with the absolute deviation on the left of the balance point?
11. Analyze the absolute deviations for the values you used in problems 3, 6, and 7. What relationship do the absolute deviations to the right of the balance point have with the absolute deviations to the left of the balance point?

**The mean for a set of data is where the absolute deviations to the right of the balance point are the same as the absolute deviations to the left of the balance point.**

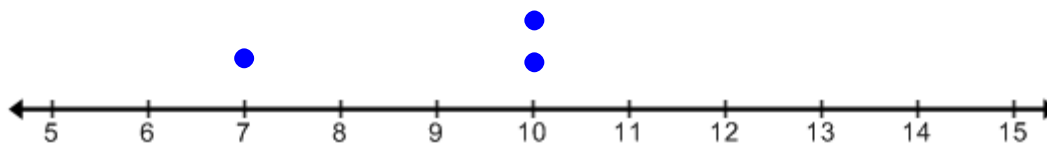
What if Carly collected handstand data from more than 2 friends? The plot below shows how long three other friends can do a handstand.



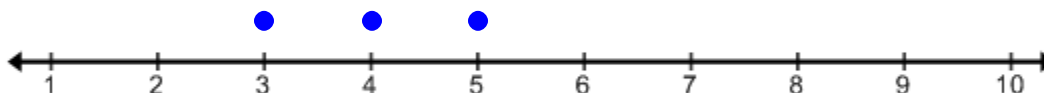
12. What is the balancing point for this data? Draw a triangle to indicate a point of balance. Use the mean and the absolute deviations of the data values to justify your answer.

13. Find a balancing point for each set of data. Draw a triangle to indicate a point of balance. Check that the sum of the absolute deviations to the left of the balance is equal to the sum of the absolute deviations to the right of the balance point. Also calculate the mean to check your answer.

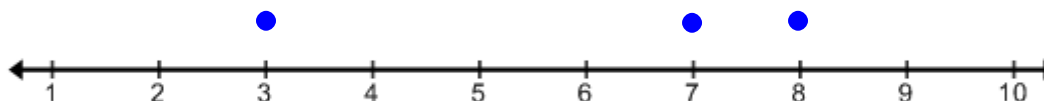
a.



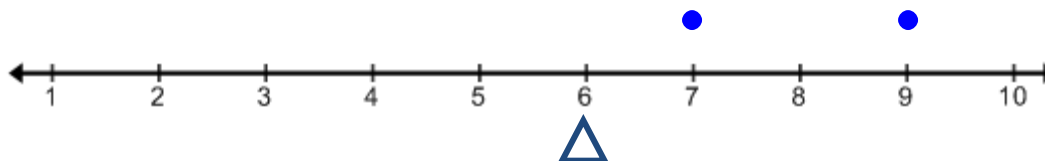
b.



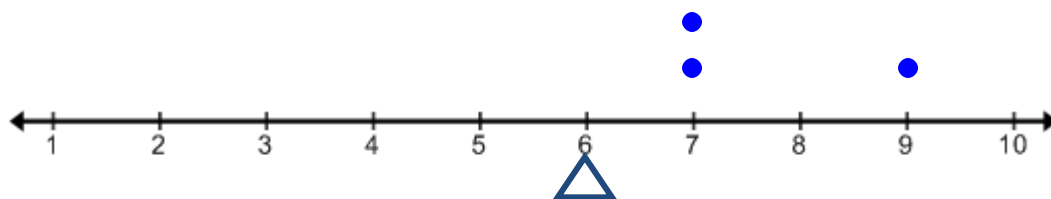
c.



14. Two values are shown in the plot below, where should a third value be placed so that the plot is balanced at 6? Check your answer by comparing the sum of the absolute deviations on each side of balance point and calculating the mean of your data points.

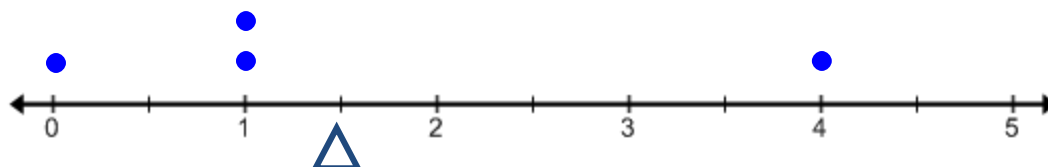


15. What if an additional value of 7 is added to the plot, where should the fourth value be placed so that the plot is still balanced at 6? Check your answer by comparing the sum of the absolute deviations on each side of the balance point and calculating the mean of your data points.



16. The dot plot shows the number of points that the Ice Rebels scored in 4 of their hockey games last season. Use the “balancing” process to explain why the mean number of goals for the team is 1.5.

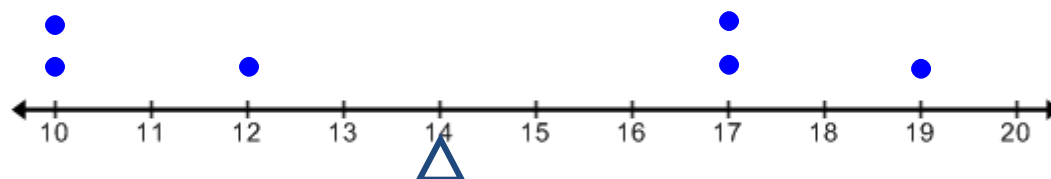
Number of Points Scored in Hockey Games



17. The weights in pounds of six loaded backpacks are shown in the plot below. The weight for the seventh backpack is missing but the mean for the weights of all seven backpacks is 14 pounds. Use the “balancing” process to find the weight of the seventh backpack.



Weights of Backpacks (pounds)



## Spiral Review

1. Write each fraction as a mixed number or whole number.

a.  $\frac{15}{4}$

b.  $\frac{43}{10}$

c.  $\frac{8}{3}$

d.  $\frac{66}{2}$

2. Solve each problem without using a calculator. Use a tape diagram if needed.

a.  $8 \times \frac{1}{2}$

b.  $30 \times \frac{1}{6}$

c.  $48 \times \frac{1}{8}$

d.  $100 \times \frac{1}{5}$

3. Estimate to the answers.

a.  $9 \times 99$

b.  $15 \times 9$

c.  $19 \times 9$

d.  $20 \times 49$

4. Find the following without a calculator.

a. One tenth more  
than five tenths.

b. One tenth more  
than twenty-five  
hundredths.

c. One hundredth  
more than thirty-  
five thousandths.

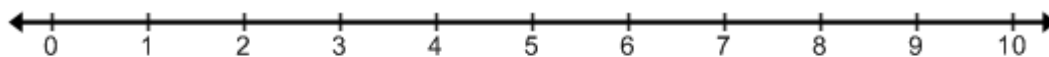
d. One thousandth  
more than 45  
hundredths.

#### 4.2c Homework: Mean as a Balancing Point

1. The number of Harry Potter books that 5 different friends have read are below.

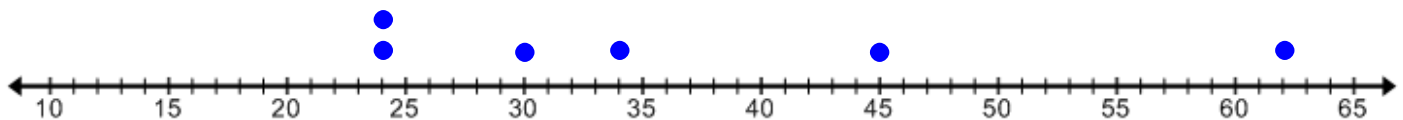
0, 5, 7, 4, 4

- a. Find the mean number of Harry Potter books read for this group of friends.
- b. Make a dot plot of the data set and check your answer by comparing the sums of the absolute deviations on each side of the mean.



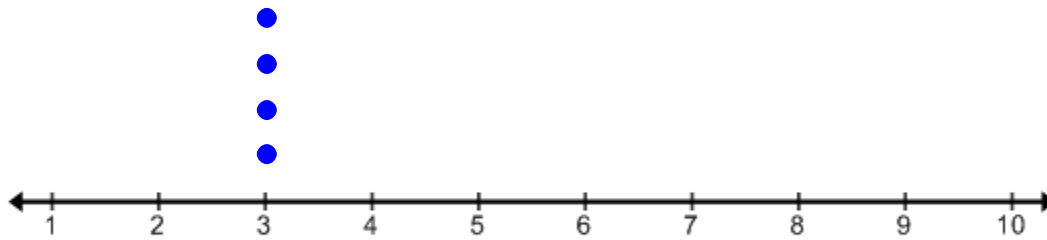
2. The number of points that the University of Utah football team scored in the first six games of their 2015 season are recorded in the plot below.

Number of Points Scored in Football Games



- a. Find the mean number of points scored for their first 6 games.
- b. Verify your answer to part a by using the “balancing” process and comparing the sums of the absolute deviations on each side of the mean.

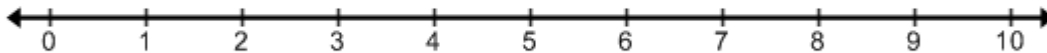
3. Use the “balancing” process to find the mean.



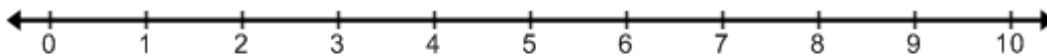
4. Use the data values given below to answer the questions that follow.

4, 5, 9, 3, 9

- a. Plot these numbers on the dot plot given below and use the “balancing” process to find the mean. Draw a triangle to represent the mean.

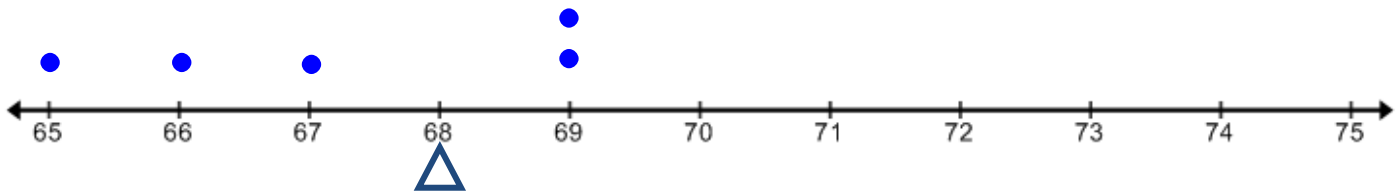


- b. Rearrange the numbers from the plot above so that the mean or balance point remains the same. Explain your process.

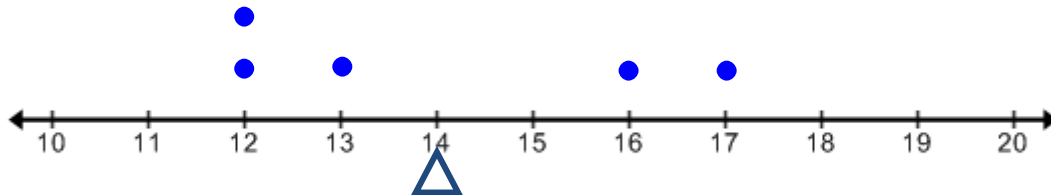


5. The weights in pounds of five male German Sheppard dogs are shown in the plot below. The weight for the sixth dog is missing but the mean for the weights of all six dogs is 68 pounds. Use the “balancing” process to find the weight of the sixth dog.

Weight of Male German Sheppard Dogs  
(pounds)



6. Find, Fix, and Justify  
Lila is examining the plot below. She is trying to decide if the balance point given is correct.



She makes the following statement.

The balance point is not correct because the numbers to the left of the balance point are 12, 12, 13, which sum to 37. The numbers on the right are 16 and 17, which sum to 33. These totals are not the same so 14 cannot be the balance point.

Lila has made a mistake. Find her mistake and justify why 14 is the balance point of the data value.



#### 4.2d Class Activity: Median and Mode

Recall that a *measure of center* describes a data set by summarizing all the values with a single number. There are other ways to measure the center of a data set in addition to finding the mean (the balancing point). These other measures of center are called the **median** and **mode**.

Directions: Once again grab a handful of linking cubes, count the number of cubes that you have, and write this number on a sticky note.

1. Explain how you found the **median** of the linking cube data below:
2. Explain how you found the **mode** of the linking cube data below:
3. What is the mean of the data set?
4. Compare and contrast the mean, median, and mode. Is there one measure of center that best represents your data?

**Median:** The median is found by ordering the data from least to greatest and the finding the middle number. If there are two numbers in the middle then the median is the mean of those two middle numbers.

**Mode:** The mode of a data set is the value or values that occur most often. There can be one mode, no mode, or more than one mode.

## Examples

Find the median and mode of each data set

- a. Data Set: 9, 10, 13, 10, 8, 11, 12

Median: Order the data from least to greatest. The median is the middle number.

8, 9, 10, 10, 11, 12, 13

The median is 10.

Mode: The mode is the value that occurs most often

9, 10, 13, 10, 8, 11, 12

The mode is 10.

- b. Data Set: 4, 7, 4, 8, 9, 5, 3, 7

Median: Order the data from least to greatest. If there is an even number of data values then the median is the mean of the two middle values.

3, 4, 4, 5, 7, 7, 8, 9



$$\frac{5+7}{2} = \frac{12}{2} = 6$$

The median is 6.

Mode: The mode is the number that occurs most often.

4, 7, 4, 8, 9, 5, 3, 7

The mode is both 4 and 7.

5. The list below shows the words per minute that a group of 12 students can type.

Words typed per minute: 40, 44, 49, 30, 40, 42, 42, 48, 60, 36, 42, 45

- a. Make a dot plot of the data.

- b. Find the mode of the data.


- c. Find the median of the data.

- d. Do you think that the median and mode will always be the same number?

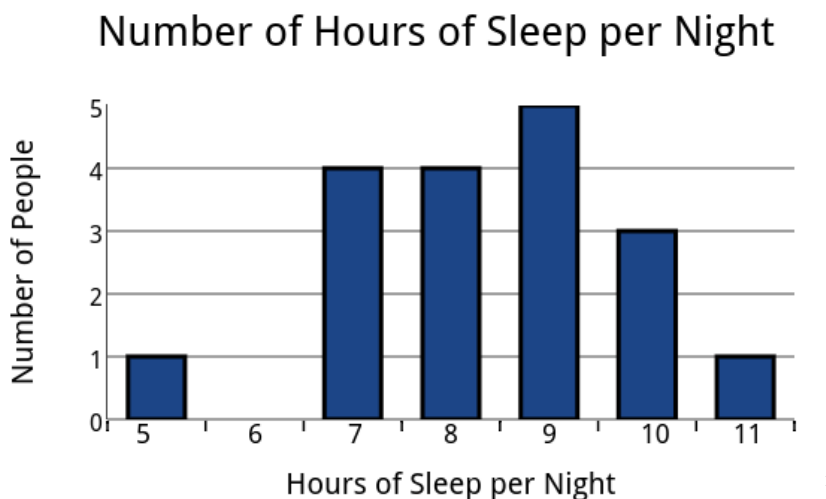
**Directions:** Find the median and mode for each set of data.



6. The table shows the number of students from each grade that are participating in the school Fun Run.

	Grade	Kindergarten	1st	2nd	3rd	4th	5th	6th
	Number of Students	75	60	55	63	40	58	65

7. The bar graph shows the average amount of sleep per night in hours for students in Ms. Kekoa's 6<sup>th</sup> grade class.

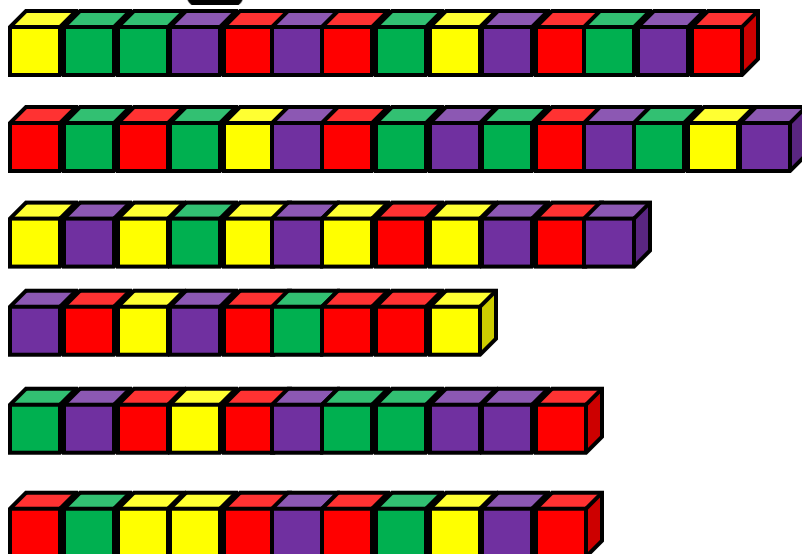


8. Amount of medicine (in milliliters) given to patients: 2.5, 7.5, 5, 5, 10, 5, 7.5, 5, 2.5, 10, 7.5, 7.5

9. Pounds of flour used per week at different bakeries: 300, 200, 250, 225, 250, 275

# 10. Find, Fix, and Justify

A group of 6 students grabbed a handful of linking cubes. Each person attached their cubes into the rods shown below.



Monte and Ava have been asked to find the median number of cubes in a handful. Their answers are given below, but they have made a mistake in their reasoning. For each person, find their mistake and fix it.

a. Monte's Answer

b. Ava's Answer

☐ 14, 15, 12, 9, 11, 11

*The median is 10.5*

☐ 9, 11, 11, 12, 14, 15

*There is no median*

c. What is the mode for the data set?

d. What is the mean number of cubes in a handful?

# 11. Find the mean, median, and mode for each set of data. Round your answer to the nearest hundredth.

a. 13.22, 11.05, 10.77, 15.04, 12.3, 12.89, 14.7, 16.3, 13.9

Mean-  
Median-  
Mode-

b.  $5\frac{1}{2}$ ,  $3\frac{3}{8}$ ,  $1\frac{1}{8}$ , 2,  $3\frac{5}{8}$ ,  $2\frac{3}{4}$ ,  $1\frac{1}{8}$

Mean-  
Median-  
Mode-

#### 4.2e Class Activity: Choosing the Best Measure of Center

1. Becky is the head of a small tech company. The salaries for all of her employees, including herself, are shown in the table below.

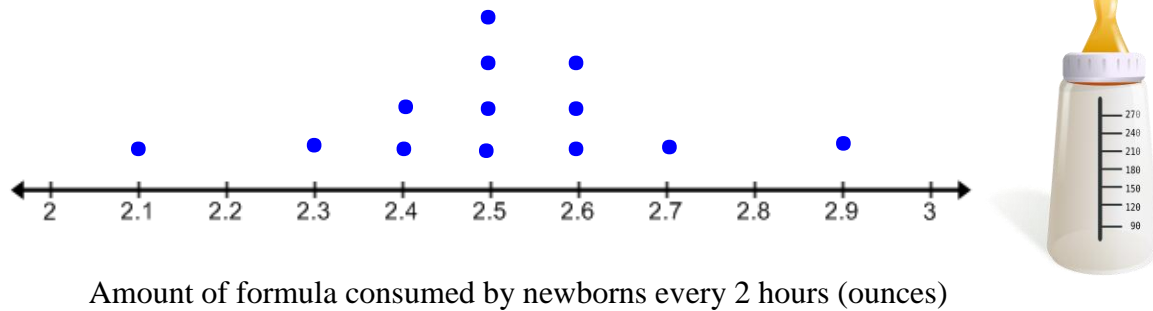
Title	Annual Salary	Number of People	Total
Chief Executive Officer	\$110,000	1	\$110,00
Marketing Manager	\$80,000	1	
Sales Manager	\$80,000	1	
Project Managers	\$50,000	2	\$100,000
Developers	\$40,000	4	
Office Assistants	\$30,000	3	
<i>Total</i>			

- How many people work at Becky's company?
- Complete the table to find the total annual salary bill for the company?
- Find the mean (average) salary for employees at this company. Round your answer to the nearest dollar.
- Julia looks at the table and claims that the mode for this data is \$80,000. What mistake has Julia made and what is the correct mode?
- What is the median salary for Becky's company?
- List all the measures of center and decide which measure of center do you think best represents the data? Justify your answer.



- g. Make a dot plot of the salary data for Becky's company and use it to justify or change your answer from part f.
- h. If Becky wants to attract future employees with a desirable average pay, which measure of center should she use and why?
- i. Becky's company does not do very well during the next year of business. She decides not to pay herself. Which measure(s) of center will not change?

2. Bethany works as a certified nursing assistant (CNA). Her job is to record the amount of milk-based formula that each newborn baby at a hospital consumes every 2 hours. The dot plot below shows this data.



- a. Find the mean, median, and mode for the data set. Round to the nearest tenth.

Mean-

Median-

Mode-

- b. Use the shape of the data to justify which measure of center best represents the data.

- c. Why might this information be useful?

3. Raul has been training for a road bike race that takes a cyclist from the town of Logan, Utah to Jackson Hole, Wyoming. Raul's times for this exact same race the previous 8 years are shown in the table below.

Raul's Bike Race Times (Hours)							
9.2	9.15	9.5	12.08	9.2	9.4	10.9	9.8

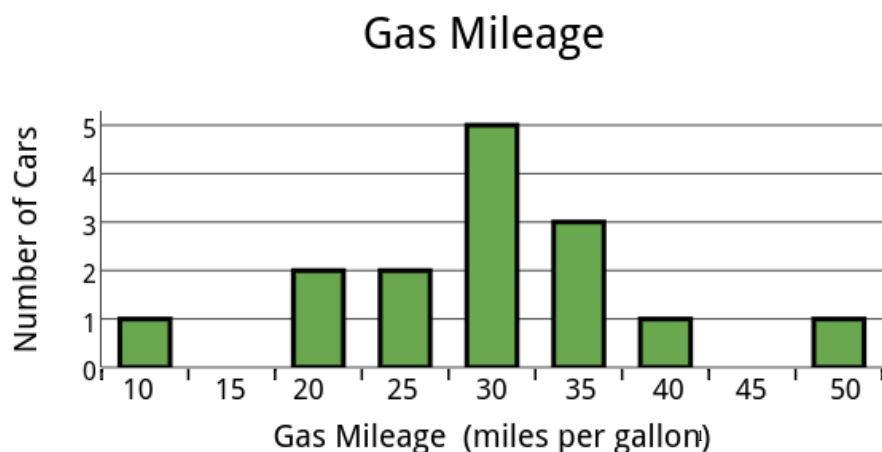
- a. Make a dot plot of this data. Discuss the shape of the data and use it to make a prediction about which measure of center will best summarize the data.



- b. What is Raul's *mean* for these previous races? What does this *mean* tell you in terms of the context? Round to the nearest hundredth.
- c. What is Raul's *median* time for these previous races? What does this *median* tell you in terms of the context?
- d. Which is the better measure of center for this data?



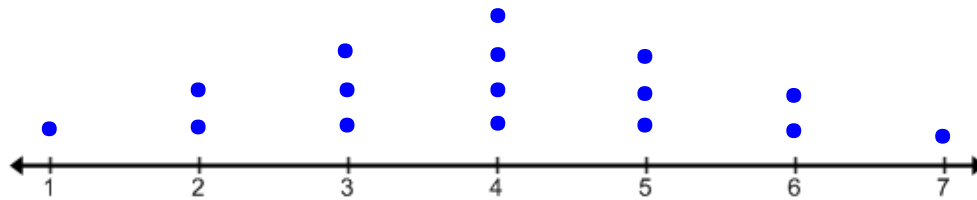
4. The bar graph shows gas mileage of several cars. Use the shape of the data to determine which measure of center best summarizes the data. Be sure to account for clusters, peaks, gaps, and outliers. Then find this measure of center and describe what it means in terms of the context.



5. What does the shape of the data tell you about which measure of center to use to summarize the data?



6. Use the dot plot below to answer the following questions.



- Find the mean, median, and mode.
- Move the least number of points so that the mean, median, and mode still remain the same value.
- Move one point so that the mean is greater than the median.
- Move one point so that the mean is less than the median.

### Spiral Review

- For each written phrase, write a numerical expression and then evaluate your expression.
  - Two thirds of the sum of forty-eight and twelve.
  - Four times the difference of three fourths and one half.
  - 5 less than one third of 21.
- Lucy and Rachel are frosting cupcakes. Lucy can frost 5 cupcakes in 1 minute and Rachel can frost 8 cupcakes in 1 minute.
  - Fill in the chart using the information given above.

Minutes	Lucy	Rachel
1	5	8
2		
3		
4		

- Identify a pattern or rule that describes how long it takes each girl to frost cupcakes.

## 4.2e Homework: Choosing the Best Measure of Center

**Directions:** For each set of data, find the mean, median, and mode. If needed, plot the data using a dot plot, bar graph, or histogram. Then justify which measure of center you think best summarizes the data. Be sure to use the shape of the data, accounting for clusters, peaks, gaps, and outliers if needed.

1. The number of siblings for each of 10 students in a 6<sup>th</sup> grade class is shown below.

2, 3, 7, 0, 1, 1, 11, 0, 1, 2

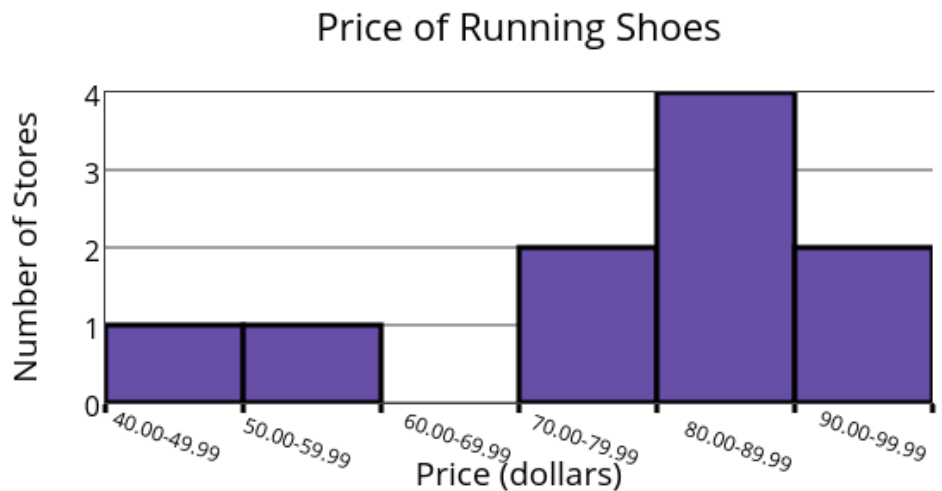
Mean:

Median:

Mode:

Which measure of center best describes the data?

2. The graph shows the price of the same running shoe at several different stores.



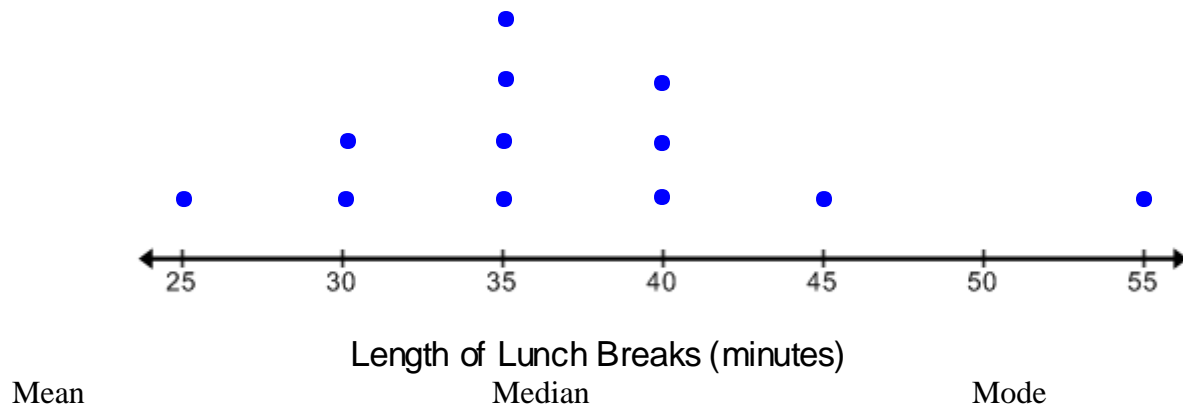
Mean:

Median:

Mode:

Which measure of center best describes the data?

3. The dot plot shows the length of lunch breaks at several different middle schools.



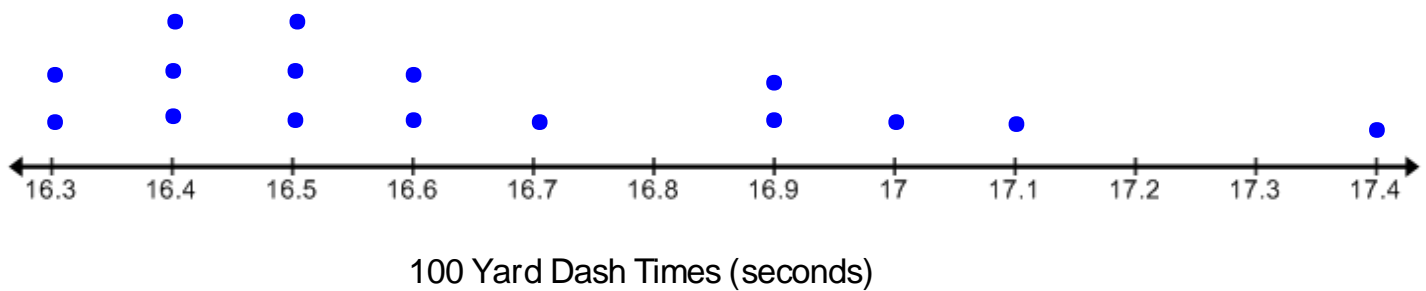
Which measure of center best describes the data?

4. In your homework assignment for **Lesson 4.1b** you collected data for a statistical question that you created and asked to several people in your class. You made a dot plot of the data and analyzed the shape of the data. Return to this data and find its mean, median, and mode. Then use the shape of the data to determine which measure of center best summarizes the data. Be sure to justify your answer.

Mean                                      Median                                      Mode

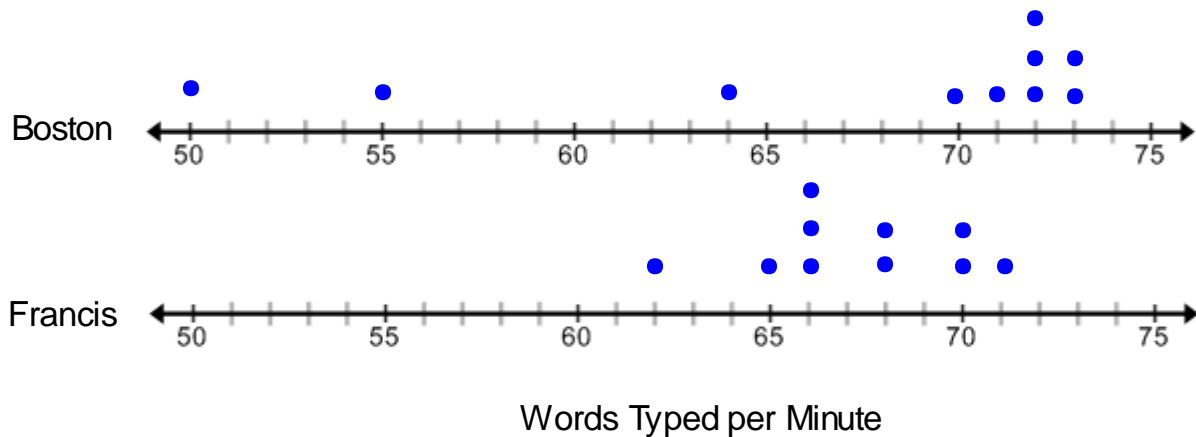
Which measure of center best describes the data?

5. Addison has just completed a 100-Yard Dash with a time of 16.6 seconds. Circle the statements below that are true.



- A. Addison's time is greater than the mode of the data.
- B. Addison's time is less than the median.
- C. Addison's time is greater than all the measures of center.
- D. Addison's time is within one-tenth of the mean.

6. Boston and Francis are practicing their keyboarding skills. They have each taken a practice test 10 times. Their scores are recorded in the dot plots below. They both have a **mean** typing score of 67.2 words per minutes but different median scores.



- a. Find the median words per minute typed for each person.
- b. To qualify to move on to the next class, you must have a final test score with a typing speed of at least 69 words per minute. You only get to take the final test once. Which person is more likely to pass the test if they take it right now? Justify your answer.

7. Use the data given to answers the questions that follow.

8, 2, 2, 5, 3, 5, 3,

- a. Find the mean and median for this set of data
- b. Suppose each data value was increased by 10. How would this change the mean, median, and mode?
- c. Suppose each data value was decreased by 2. How would this change the mean, median, and mode?
- d. Find one data value that if added to this set of data will keep the median the same.

#### 4.2f Self Assessment: Section 3.2

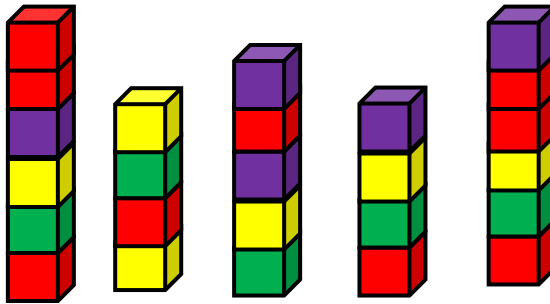
Consider the following skills/concepts. Rate your comfort level with each skill/concept by checking the box that best describes your progress in mastering each skill/concept. Corresponding sample problems, referenced in brackets, can be found on the following page.

Skill/Concept	Minimal Understanding 1	Partial Understanding 2	Sufficient Mastery 3	Substantial Mastery 4
1. Understand the mean as a “leveling out” of data or a fair share. Also understand how the mean can be interpreted as the “balancing point” for the data values.	I can kind of explain how the mean describes a fair share of the data. I do not know how it can be interpreted as the “balancing point”.	I can explain how the mean describes a fair share of the data. I can find the absolute deviations of the values in the data set but I do not know how they relate to the mean.	I can explain how the mean describes a fair share of the data. I struggle to relate this to the formula for the mean. I can show how the mean is the “balancing point” of the data by comparing the sum of the absolute deviations of the values to the right and left of the mean.	I can explain how the mean describes a fair share of the data and relate it to the formula for finding the mean. I can show how the mean is the “balancing point” of the data by comparing the sum of the absolute deviations of the values to the right and left of the mean.
2. Find the mean from a set of data and interpret its meaning for a given context.	I can find the mean from a set of data but don’t know what it means if the <i>mean</i> is not a whole number.	I can accurately find the mean from a set of data that is presented as a list.	I can accurately find the mean from a set of data that is represented in a variety of ways. I can interpret what the <i>mean</i> means for a given context.	I can accurately find the mean from a set of data that is represented in a variety of ways. I can interpret what the <i>mean</i> means for a given context and give suggestions for how the values were collected.
3. Find the median and mode from a set of data and interpret their meaning for a given context.	I know what the median and mode are but do not know how to find them.	I know how to find the median if there is an odd number of data values. I do not know how to find the median if there is an even number of data values. I can find the mode if there is only one.	I can find the median and mode from a set of data that is presented in a list. I understand that there can be more than one mode or no mode at all. I can interpret what the median and mode mean for a given context.	I can find the median and mode for any set of data that is represented in a variety of ways. I understand that there can be more than one mode or no mode at all. I can interpret what the median and mode mean for a given context and given suggestions for how the values were collected.
4. Determine the best measure of center for a given data set and justify why it is the best measure of center.	I know that there are different measures of center but don’t really know how to choose one that best represents the data.	I can sometimes guess which measure of center is best but don’t really know how to determine which measures it best for any given set of data.	Given a set of data, I can choose the measure of center that I think best represents the data. I struggle to construct good arguments and give justification as to why I think the measure of center best represents the data.	Given a set of data, I can choose the measure of center that I think best represents the data. I can give justification and construct good arguments as to why I think the measure of center best represents the data.

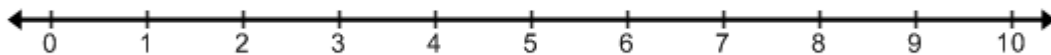
## Sample Problems for Section 4.2

*Square brackets indicate which concept/skill the problem aligns to.*

1. A group of five students have each grabbed a handful and linking cubes. They each stick their cubes together into a rod. These rods are shown below. [1]



- a. Show or explain how you can equally redistribute the blocks so that everyone gets the same number of blocks.
- b. Find the *mean* number of blocks there are per handful using the mean formula and explain how this relates to the fair share or unit rate.
- c. Verify your *mean* by plotting each number of blocks in a handful on a number line and showing that the *mean* is the balancing point. Be sure to discuss how you know it is balanced with absolute deviations.

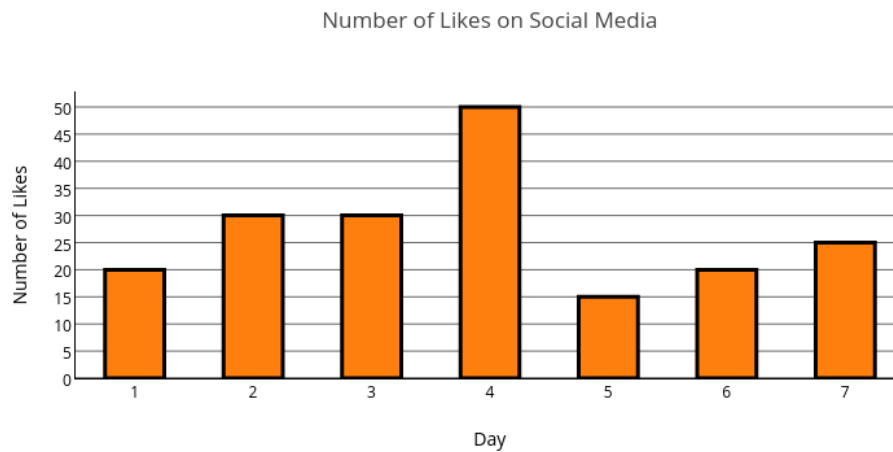


2. Find and interpret the mean, median, and mode for each set of data below. Then determine which measure of center best represents the data. Be sure to justify your answer. [2][3][4]

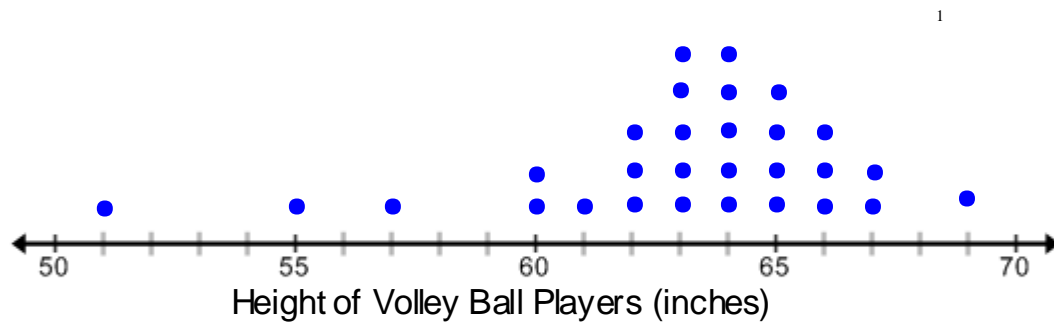
a.

Number of sit ups			
78	86	86	95
90	71	110	102
92	80	106	100

b.



c.





## Section 4.3: Measures of Variability

### Section Overview:

At this point in the chapter, student's ability to interpret and analyze data should be growing. In addition to using shape and center to describe a data set, they can use variability. In this section, students will learn how variability is an important component to fully understanding a data distribution. Finding the range is one way that students can discuss the variability of a data set. However, it does not tell us what is happening with the values within the data set itself. For distributions where the mean is a better measure of center, variability is commonly measured in terms of how far the values deviate from the mean. Students will use the work they did in the previous section with absolute deviations to calculate the mean absolute deviation. This will set the stage for work that they will do in later years with standard deviation. In cases where the median is a better measure of center, students will use the range and interquartile range (IQR) to describe the measure of variability. The IQR is a numerical measure that gives the spread of the middle 50% of the data points. Finally, students will learn how to construct and analyze a box plot which gives a snap shot of the 5-number summary and can help students to visualize the shape of the data. The numerical measures of spread and box plots aid in the comparison of data as students will be able to quickly see how two or more data sets relate to one another.

### Primary Concepts and Skills to Master in this Section:

*By the end of this section, students should be able to:*

1. Find and interpret the range for a data set.
2. Find and interpret the mean absolute deviation (MAD) for a data set.
3. Find and interpret values of the 5-number summary for a data set.
4. Find the interquartile range (IQR) for a data set and interpret its meaning.
5. Display numerical data in a box plot, use key terms to describe its shape, and interpret the distribution of data.
6. Draw conclusions about two or more data sets by comparing their box plots.

### 4.3a Class Activity: Variability and Mean Absolute Deviation

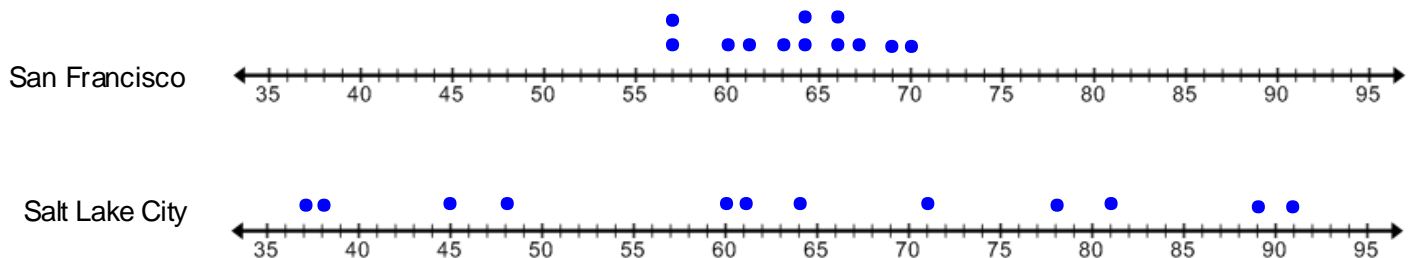


1. Olivia has been offered a job in both Salt Lake City and San Francisco. She is trying to decide which city she would rather live in and climate is very important to her. She has heard friends that live in Utah talk about the great skiing and friends that live in California talk about the nice warm beaches. She decides to look at the average monthly temperature in degrees Fahrenheit for each state.

City	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
Salt Lake City	37°	45°	60°	61°	71°	82°	91°	89°	78°	64°	48°	38°
San Francisco	57°	60°	61°	63°	64°	66°	66°	67°	70°	69°	64°	57°

- a. How would you describe the temperatures for each city?
- b. Find the mean temperature for each city. Round your answer to the nearest degree.
- c. Olivia is surprised that the mean monthly temperature for both cities is the same. Using the mean temperature as the only factor, where would you tell Olivia to move to.
- d. Are there any other factors relating to the monthly temperatures that can help Olivia make her decision? Use the dot plots below to answer this question.

Average Monthly Temperatures  
in Degrees Farenheit



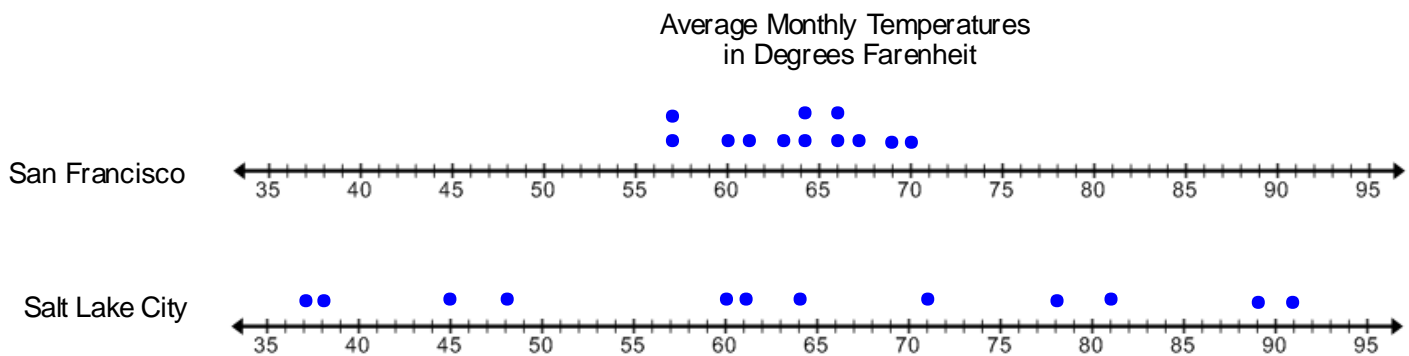
The *range* of a data set is the difference between the greatest value (maximum) and the smallest value (minimum).

$$\text{Range} = \text{Max} - \text{Min}$$

- e. Find the range for each city.

Range is not the only measure of variability. **Variability** is used in statistics to tell us how spread out data is from some focal point. We can also look at how spread out the data is from a measure of center, in this example the mean.

- f. Mark the mean on both of the dot plots with a balance triangle. Describe the variability of the monthly temperatures to their mean for each city; determine which city has more variability in its average monthly temperatures.



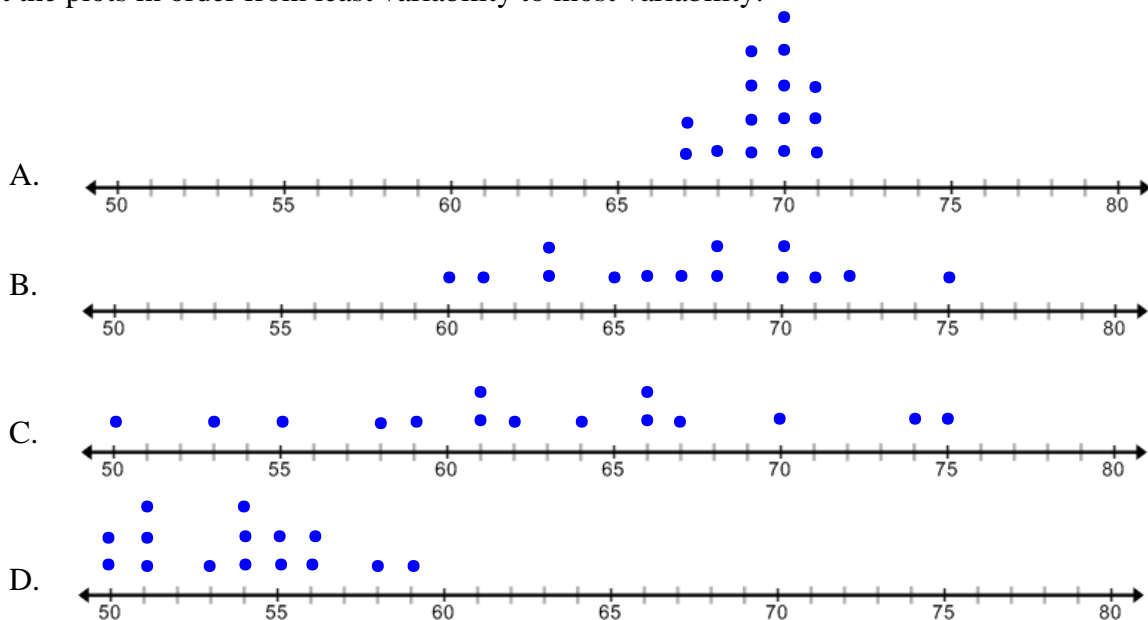
- g. How does determining the level of variability in the data help Olivia make her moving decision?

- h. For which data set is the mean a better indicator of a typical value and why?

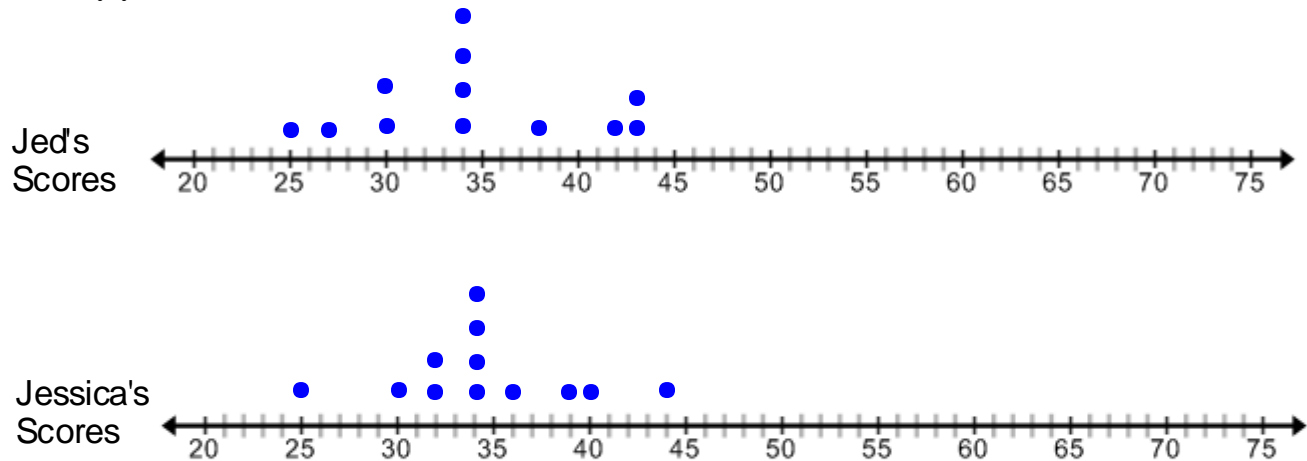
2. A coach is trying to decide which softball pitcher he should play for their next game. He is looking at the numbers of runs that each of his two pitchers have allowed in their last 10 games. The data for each pitcher is shown on the dot plot below.



- Find the mean number of allowed runs for each pitcher.
  - For which distribution does the mean give a better indicator of a typical value? Explain your answer.
  - The softball coach states that both pitchers have the same mean number of runs allowed for their previous ten games so it does not matter which girl he should play. Is he correct in his thinking? Explain.
3. List the plots in order from least variability to most variability.



4. The plots below show the scores for two skateboarders from their last 12 skate boarding competitions. The mean score for both athletes is **34.5**. Which athlete appears to have less variability in their scores? Justify your answer.



In the problem above both plots shows similar variability. Recall that variability shows us how spread out the data is from some focal point, in this case the mean. In previous lessons, we used **absolute deviations** to analyze how far away each data point is from the mean.

5. Use what you know about **absolute deviations** to try and find a better way to measure the variability for Jed's data.

- Find the absolute deviations for each data value from the mean for Jed's data. Record them above each value on the dot plot.
- Find the sum of the absolute deviations and divide by the total number of observations.

We now have one number for Jed's scores that represents the average deviation of the data points from the mean. This is called the **mean absolute deviation (MAD)**.

- Find the mean absolute deviation for Jessica's data.

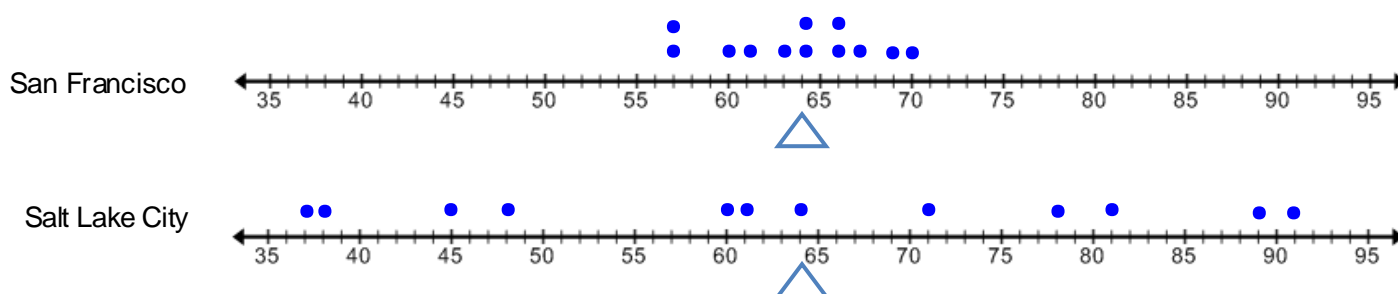
d. What does the mean absolute deviation (MAD) for each skate boarder mean?

e. Which athlete has a higher MAD? What does this mean?

**Mean Absolute Deviation, or MAD, is a measure of variation. It is computed by finding the mean of the absolute deviations in the distribution.**

6. Find the Mean Absolute Deviation for the Salt Lake City and San Francisco Temperature Data given earlier. The dot plots for both cities are below, recall that the mean temperature for both cities is  $64^{\circ}$ . Use the tables given if needed.

Average Monthly Temperatures  
in Degrees Farenheit



San Francisco Average Monthly Temperatures

Temperature	Expression that shows counting distance	Absolute deviations
57		
57		
60		
61		
63		
64		
64		
66		
66		
67		
69		
70		

Sum of the Absolute Deviations:  
MAD:

Salt Lake City Average Monthly Temperatures

Temperature	Expression that shows counting distance	Absolute deviations
37		
38		
45		
48		
60		
61		
64		
71		
78		
81		
89		
91		

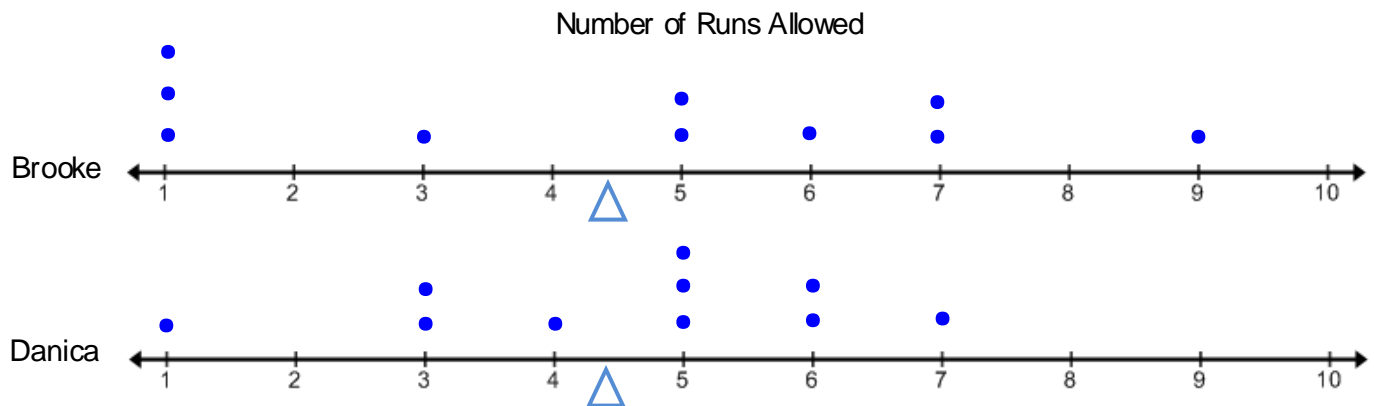
Sum of the Absolute Deviations:  
MAD:

7. What does the mean absolute deviation for each set of data tell us?

8. Fill in the blanks in the statement below with the words “greater” or “smaller”.

The mean absolute deviation describes how each data value varies from the mean. The higher the MAD the \_\_\_\_\_ the variability, the lower the MAD the \_\_\_\_\_ the variability.

9. Find the mean absolute deviation (MAD) for Brooke and Danica’s softball data. Recall the *mean* number of runs allowed for both pitchers is 4.5.



Brooke’s MAD:

Danica’s MAD:

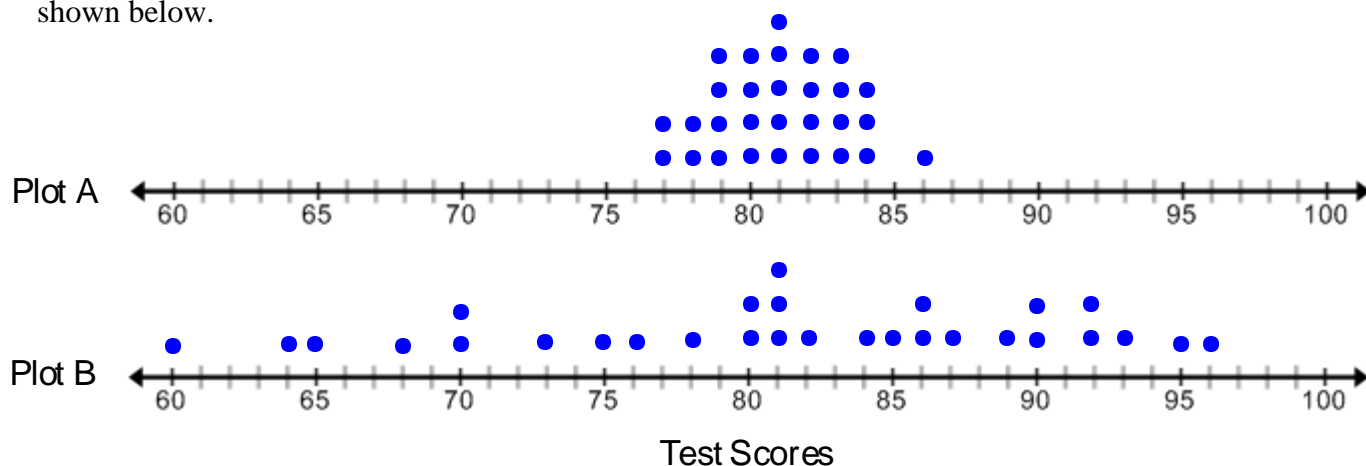
10. Use the data given below to answer the following questions.

60, 60, 60, 60, 60, 60, 60

- a. Plot the data on a dot plot
- b. Find the mean and the MAD of the data. Interpret what the MAD tells us about the amount of variability in the data.
- c. Suppose an additional data value of 61 is added to this data set. Add this value to the dot plot. Without doing any calculations determine how this will affect the mean? How will this affect the MAD?
- d. Now suppose that in addition to adding 61 to the data set a value of 59 is added as well. Without doing any calculations determine how this will affect the mean? How will it affect the MAD?
- e. Suppose an additional value was added at 65. What single value must also be added to the plot to keep the mean at 60? Explain your answer.
- f. Suppose an additional value was added at 64. What two additional values must be added to the plot to keep the mean at 60? Explain your answer.



11. Raphael has just received his test scores for English and Science. The bell has just rung to go home and he is in too much of a hurry to write down his scores. His teacher gives him a paper on the way out the door that shows a dot plot of the class scores (out of 100) for each test. The dot plots are not labeled but Raphael's teacher told the class that both tests had the same average score of 81. The dot plots are shown below.



- Find the range of data for each plot.  
 Plot A Range:  
 Plot B Range:
- Without doing any calculations, which dot plot represents a class with a higher MAD? Explain.
- On the way home from school Raphael is thinking about his tests scores and can't remember them. He does remember that his teacher told him that he received one of the high scores in the class for his English test. Which dot plot would Raphael want to be the distribution of scores for his English test if he had one of the high scores? Justify your reasoning.
- He also remembers that his score on his science test was below average. Which dot plot would Raphael want to be the distribution of scores for his science class? Justify your reasoning.

12. Claudia loves fresh peppers from the garden. She has decided to plant two varieties of bell peppers, Red Bellas and Beauty Bells. For each variety, she plants 5 plants and grows them in the same conditions. The table below shows the number of peppers that each of pepper plants produced.

Plant	Red Bellas		Beauty Bells	
1	7		14	
2	8		10	
3	10		5	
4	8		6	
5	9		16	



- a. Draw a dot plot to represent the number of peppers produced for each plant.
- b. Find the *mean* number of peppers produced for each variety of pepper.
- Red Bellas *mean*:
- Beauty Bells *mean*:
- c. Find the MAD for each variety of pepper. Use the table or the dot plot to track the deviations.
- Red Bellas MAD:
- Beauty Bells MAD:
- d. Claudia is only going to plant one kind of pepper plant next season and she only cares about the number of peppers a plant will produce. What pepper plant should she choose next season? Use the *mean* and MAD for each variety of pepper to justify your answer.

## Spiral Review

### Multiplying Decimals

1.  $5 \times 3.21$

2.  $8 \times 0.001$

3.  $3.012 \times 3$

4.  $3.2375 \times 5$

### Add and Subtract Decimals

5.  $0.8 + 0.007 - 0.004$

6.  $5 + 0.682 - .03$

7. Four plus seventy-  
three thousandths  
minus four hundredths

8. Four and three  
tenths minus five  
hundredths.

### Measurement Conversions

9.  $1 \text{ ft} = \underline{\hspace{1cm}} \text{ in}$

10.  $2.5 \text{ ft} = \underline{\hspace{1cm}} \text{ in}$

11.  $54 \text{ in} = \underline{\hspace{1cm}} \text{ ft}$

12.  $36 \text{ in} = \underline{\hspace{1cm}} \text{ yd}$

### Finding Volume

13. Find the volume of a rectangular prism that has a length of 7 inches, a width of 5 inches, and a height of 9 inches.

### 4.3a Homework: Variability and Mean Absolute Deviation

1. You are trying to decide what brand of trail mix to buy. You think that the best trail mix has a lot of M&Ms in it. The table below shows the amount of M&Ms in 7 randomly selected bags of each brand of trail mix.

Trail Mix Bag	1	2	3	4	5	6	7
Hiker's Delight	30	27	34	29	30	31	29
Salty and Sweet Mix	36	40	21	27	31	25	30

- a. What is the range of the number of M&M's for each brand of trail mix?

Hiker's Delight range:

Salty and Sweet Mix range:

- b. Make a dot plot for each set of data. Be sure to use the same scale for each plot so that you can compare the distributions.

- c. Find the *mean* number of M&M's for each brand of trail mix.

Hiker's Delight *mean*:

Salty and Sweet Mix *mean*:

- d. Use the table below to find the MAD for each set of data. Explain what the MAD means for this context.

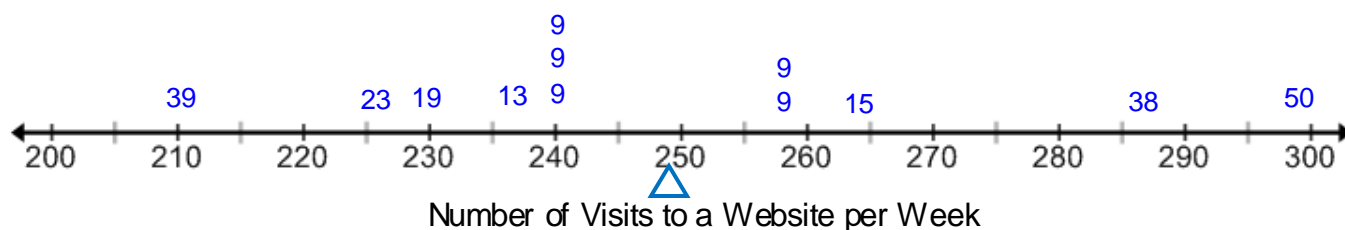
Hiker's Delight	30	27	34	29	30	31	29
Absolute deviation	0	3	4	1	0	1	1
Salty and Sweet Mix	36	40	21	27	31	25	30
Absolute deviation	6	10	9	3	1	5	0

Hiker's Delight MAD:

Salty and Sweet Mix MAD:

- e. For which distribution is the mean a better indicator of a typical value? Explain your answer.
- f. Considering variability, which brand of trail mix do you prefer? Explain your answer.

2. In the dot plot given below, each value is shown as its absolute deviation from the *mean* of **249**.

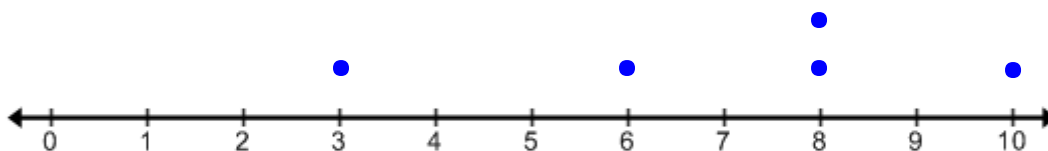


- a. Find the MAD of the data.



- b. Which of the following is the correct interpretation of the MAD for this context?
  - i. The number of visits to the website per week is approximately 20.2 visits away from the mean.
  - ii. The number of visits to the website per week is approximately 20.2.
  - iii. The number of visits to the website per week is, on average, approximately 20.2 visits from the *mean* of 249.

3. Use the data given below to answer the following questions.



- a. Find the *mean* of the data and mark it with a balance triangle on the plot.
- b. Mark the absolute deviations from the *mean* for each data value on the dot plot and use them to find the MAD.

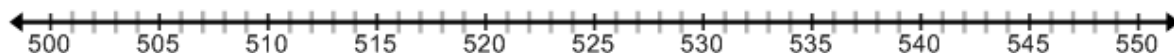
c. What value could you add to this data set that will not change the *mean*? Explain your answer.

d. Suppose a value of 1 is added to the data set. Add this point to the dot plot. Without doing any calculations determine how this will affect the mean and MAD.

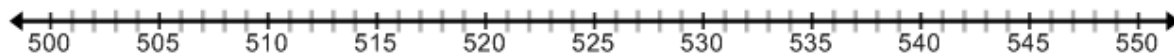
e. What value or values would you add to the data set in addition to the value of 1 to make the *mean* return to 7?

4. For the number lines below, make a dot plot with at least 6 values that matches the following variability requirements.

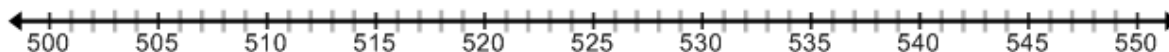
a. A plot with a lot of variability.



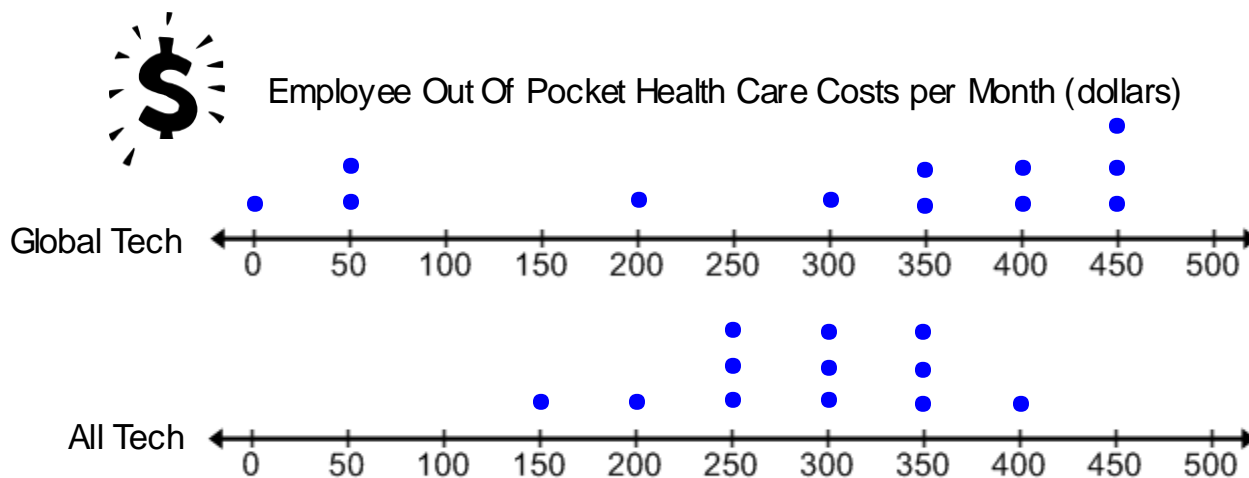
b. A plot with less variability than the plot in part a.



c. A plot with no variability.



5. The dot plots below show the out of pocket cost of health care per month for several employees at two different technology companies. The *mean* cost of health care for both companies is approximately \$287.50.



- Describe the shape of the data for each company.
- The MAD for each company is listed below. Interpret what the MAD is telling us for these data sets.  
Global Tech MAD:  $\approx 141.7$   
All Tech MAD:  $\approx 56.25$
- For which company is the mean a better indicator of a typical value of the data distribution? Use the shape and MAD to justify your answer.
- What might be a more appropriate measure of center for the data set with the larger MAD?

6. Adila is going to purchase her first cell phone and is trying to decide between two different brands of phones. She is going to base her final decision on how long the phone can last without having to be recharged. She decides to survey her friends that own these two phones on the battery life of their phones. The results of her survey are shown in the table below.

Phone Brand A

Battery Life (hours)	Absolute Deviation
10	
9	
8	
9	
10	
7	
11	
10	
Total:	

Phone Brand B

Battery Life (hours)	Absolute Deviation
7	
14	
13	
8	
6	
8	
9	
9	
Total:	

- a. Find the mean battery life for each phone brand.

Brand A *mean*:

Brand B *mean*:

- b. Fill in the missing information in the table to find the MAD for each set of data. Explain what the MAD means for each data set as well.

Brand A MAD:

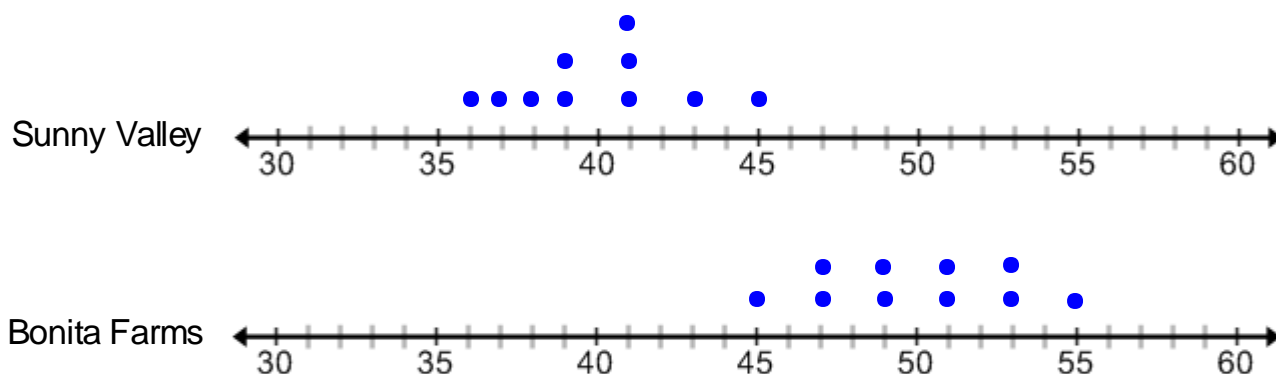
Brand B MAD:.

- c. Make an argument for choosing each brand of phone, be sure to discuss the data's variability in your argument.



7. The manager of a cafeteria is ordering potatoes. He can choose between two different brands of potatoes that are of the same quality and have the same cost per pound. Sunny Valley offers 40 lb bags and Bonita Farms offers 50 lb bags. Each manufacturer guarantees that their bags of potatoes will weigh close to the labeled weight. The manager at the cafeteria has weighed the last ten bags of potatoes that he received from each brand. His results are shown below. None of the bags are the exact weight.

Actual Weight of Bags of Potatoes (pounds)



The manager at the cafeteria is trying to determine which brand of potatoes offers bags that are closer to their labeled weight so he can decide which brand to buy in the future. Circle the statements below that are true.

- The MAD for Sunny Valley is 40 lbs.
- The MAD for Bonita Farms is 2.6 lbs.
- The MAD for Sunny Valley is 2.2 lbs.
- The MAD for Bonita Farms is 50 lbs.
- The manager cannot determine which bags are closer to their labeled weight because the average for each data set is different.
- Sunny Valley offers bags that are closer to their labeled weight because the MAD for its data is smaller than the MAD for Bonita Farms.
- The manger should go with the 50 lbs bags over the 40 lb bags regardless of their exact weights because he is always going to get more potatoes for the same price.
- The average deviation of weights from the *mean* weight of 40 pounds for Sunny Valley is smaller than the average deviation of weights from the mean weight of 50 lbs for Bonita Farms.



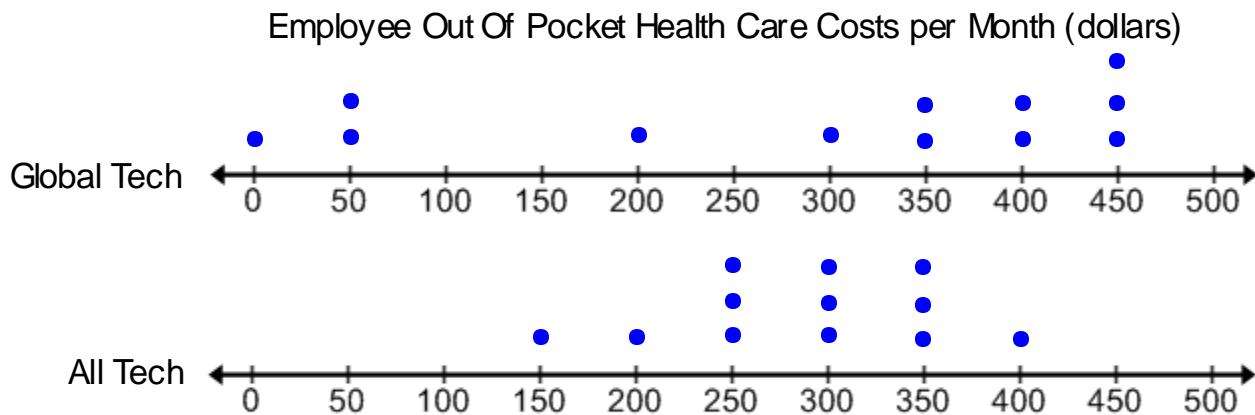
8. Two classes took the same test. A summary of the distribution of scores is provided below.

Class	Mean	MAD
1 <sup>st</sup> period	76	4
2 <sup>nd</sup> period	76	10

- Suppose you had the highest score in the class on this test. Would your score most likely have been higher if you were in 1<sup>st</sup> period or 2<sup>nd</sup> period? Explain your answer.
- Suppose you had the lowest score in the class on this test. Would your score have most likely have been higher if you were in 1<sup>st</sup> period or 2<sup>nd</sup> period? Explain your answer.

### 4.3b Class Activity: The Interquartile Range

In the previous homework assignment, you were given the out of pocket cost of health care for two different companies. This data is shown below.



1. Discuss which measure of center best describes each set of data? Justify your answer using the questions and answers from your homework assignment.
2. Find the median from the Global Tech data and compare it to the mean.
3. Discuss how to analyze the variability of a data set when the median is the best measure of center?
4. Find and interpret the range of the Global Tech data.

The range for this data set gives us an idea of how spread out the data is from min to max. However, it does not really let us know how the numbers are organized within the range. **Quartiles** can help us to better interpret the variability within the data set. Just like the median, quartiles split the data into equal parts.

- ***Quartiles* are the values of the points that split the data into four equal parts.**
- **The  $1^{st}$  quartile (or *lower quartile*) separates the lower half of the data into two equal parts. It is the median of the lower half of the data.**
- **The  $3^{rd}$  quartile (or *upper quartile*) separates the upper half of the data into two equal parts. It is the median of the upper half of the data.**

5. Find and label the median,  $1^{st}$  quartile,  $3^{rd}$  quartile, max, and min for the Global Tech data below.

0, 50, 50, 200, 300, 350, 350, 400, 400, 450, 450, 450

- a. How much of the data is below the median?
- b. How much of the data is below the  $1^{st}$  quartile?
- c. How much of the data is between the median and  $3^{rd}$  quartile?
- d. How much of the data is between the  $1^{st}$  quartile and  $3^{rd}$  quartile?
- e. How much of the data is below the  $3^{rd}$  quartile?

The **interquartile range (IQR)** can be found by finding the difference between the  $3^{rd}$  quartile and  $1^{st}$  quartile.

8. Find the interquartile range (IQR) of the Global Tech data. Then mark it on the data above.
9. How much of the data falls within the interquartile range (IQR)?
10. What does the IQR tell you about the data?

The interquartile range or *IQR* is the difference of the 3<sup>rd</sup> quartile and 1<sup>st</sup> quartile.

$$IQR = Q3 - Q1$$

The *IQR* is a single number that describes that variability of a data set.

11. Find the interquartile range of the data given below. As you work through the data write down your steps for finding the IQR.

Top Speeds of the World's Fastest Motorcycles in Miles Per Hour



186, 174, 175, 176, 176, 176, 175, 227, 190, 230, 248, 275, 169, 325, 190

12. What does the IQR tell us about this data?
13. The data sets given below show the number of words typed per minute for Boston and Francis.
- Boston: 70, 71, 50, 73, 73, 72, 55, 72, 64, 72
- Francis: 71, 70, 65, 66, 68, 70, 62, 66, 68, 66
- Find the range for the words typed per minute for each person.
  - Find the median words typed per minute for each person. Show your work above.
  - Find Q1 and Q3 for each set of data.
  - Find the IQR for each set of data.

e. Use the IQR to compare the variability of each data set.

f. Use the information above to make an argument for who will type more words per minute in the future.



14. The tables below represent the ages of people at two different city skate parks on a given afternoon.

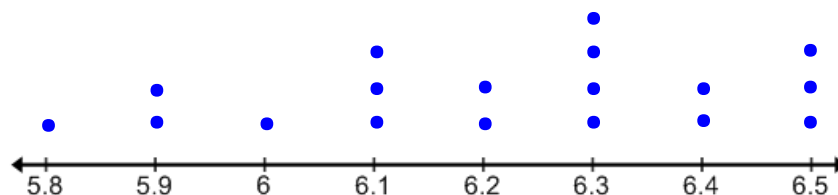
Ages of People at City Skate Parks	
South Jordan City	Salt Lake City
4	3
10	8
10	10
11	11
14	11
15	19
15	22
15	23
16	28
20	29
32	32
40	39



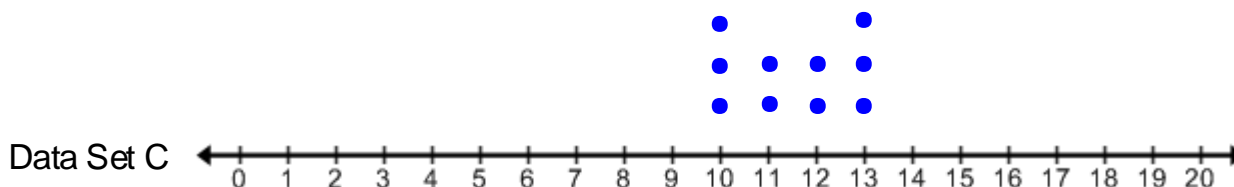
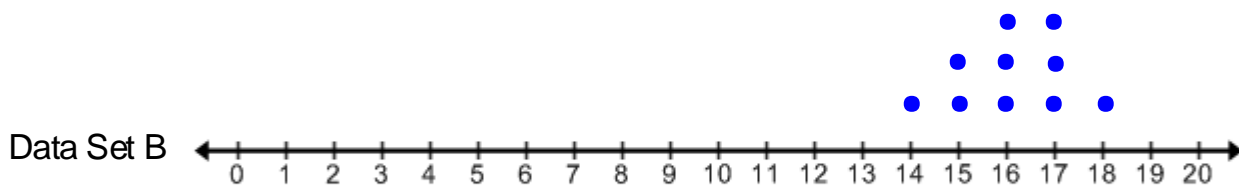
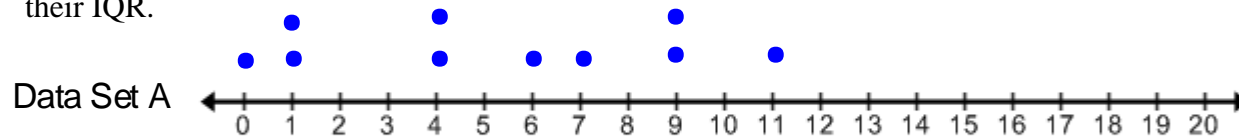
a. Make a prediction about how the sizes of IQR for the ages of people at each skate park compare.

b. Find and interpret the IQR for each city and compare your answers to your prediction in part a.

15. Find the range and interquartile range of the data given in the dot plot below.



16. Without doing any calculations order the data sets below from least to greatest according to the value of their IQR.



17. Create a data set with at least 8 data values that match each set of given conditions.

a. A data set with an IQR that is very close to the range.

b. A data set with a small IQR and a large range.

c. A data set where the median and the 3<sup>rd</sup> quartile are the same number.

### Spiral Review

Fill in each blank with  $<$ ,  $>$ , or  $=$ .

1.  $14.87$  \_\_\_\_  $14.85$

2.  $329.4$  \_\_\_\_  $3.294$

3.  $0.68$  \_\_\_\_  $\frac{68}{100}$

4.  $0.4003$  \_\_\_\_  $4.03$

Round each number to the nearest tenth.

5.  $25.43$

6.  $5.48$

7.  $2.72$

8.  $0.446$

Fill in each blank with the correct measurement conversion.

9.  $5\text{ m} = \text{____cm}$

10.  $30\text{ m} = \text{____km}$

11.  $2000\text{ mm} = \text{____cm}$

12.  $40\text{ km} = \text{____m}$

Find each product.

13.  $\frac{1}{2} \times \frac{1}{3}$

14.  $\frac{1}{2} \times \frac{1}{4}$

15.  $\frac{3}{4} \times \frac{1}{3}$

16.  $\frac{2}{5} \times \frac{2}{3}$

### 4.3b Homework: The Interquartile Range

1. Find the range for each situation given.
  - a. The youngest person in a family is 3 years old. The oldest person in the same family is 38 years old.
  - b. The longest snake in a reptile exhibit is 10.5 ft long. The shortest snake in the exhibit is 0.5 ft long.
2. The data below represents the number of points scored by a team in their first season of basketball games.
  - a. Place the words below in the correct blank in the data set.

Max

Min

Median

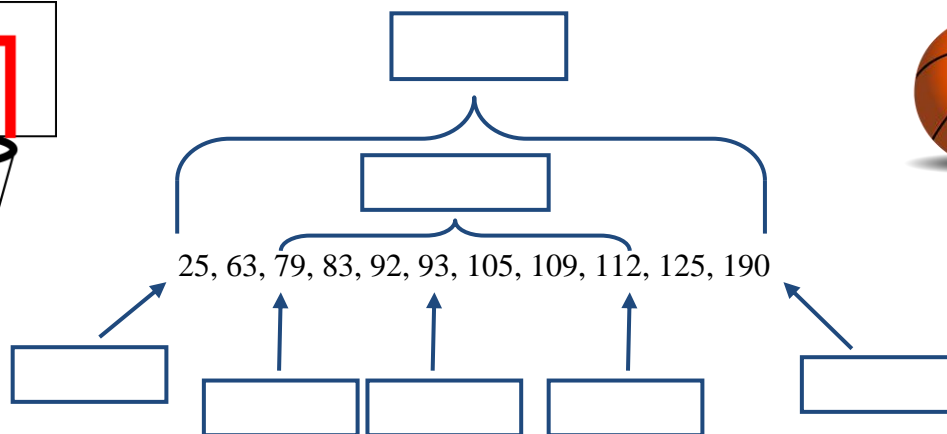
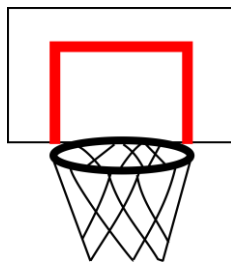
Q1

Q3

Range

IQR

Number of Points Scored in Basketball Games

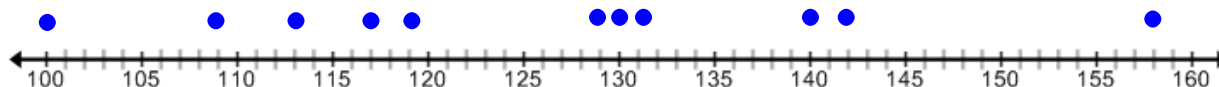


- a. Calculate the Range for the data above.
- b. In how many games did the team score more than 93 points?
- c. In 75% of the games, the team scored less than how many points?
- d. Calculate and interpret the IQR for the data above.
- e. Which measure of variability, the range or the IQR, better reflects this distribution of data? Explain.

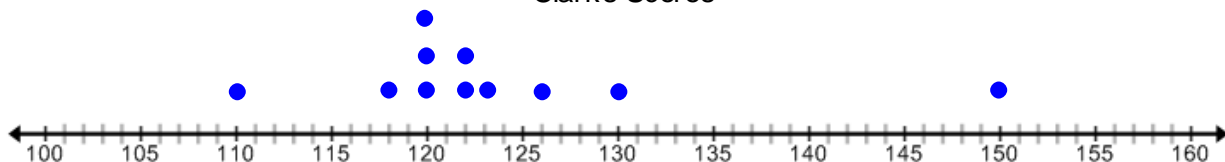


3. Clark and Savannah love to go bowling. They have recorded their bowling scores for the last 11 games they have played. These scores are shown in the dot plots below.

Savannah's Scores



Clark's Scores



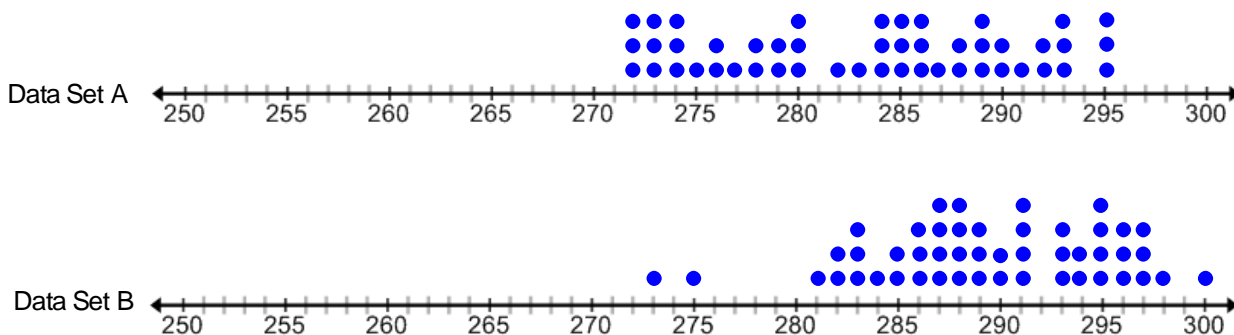
- Find the range for each person's set of scores.
- Find the median score for each person. Show your work above.
- Find Q1 and Q3 for each set of data.
- Find the IQR for each set of data.
- Use the IQR to compare the variability of each data set.
- Use the information above to make an argument for who will get a higher score in the future.



4. The lunch menu at Anthony's school advertises that each lunch comes with a serving of 12 tater tots. Anthony is convinced that the lunch server with the red hair at his school cafeteria hands out about 12 tater tots each day. His friend thinks the lunch server with the glasses hands out about 12 tater tots each day. They decide to track the number of tater tots they get on their lunch trays for two weeks. Each day Anthony goes to the lunch server with the red hair and his friend goes to the lunch server with the glasses. At the end of the two weeks they find the median number of tater tots given out by each server. Their findings are in the table below.

Number of Tater Tots										
Lunch Server	Mon	Tue	Wed	Thur	Fri	Mon	Tue	Wed	Thur	Fri
Red Hair	13	9	12	11	14	15	11	12	10	14
Glasses	15	13	11	10	13	5	20	9	17	10

- Find the median for each lunch server.
  - Find the IQR for each lunch server.
  - Use median and IQR for each lunch server to determine who is most likely to give out the advertised 12 tater tots.
5. Jonas was asked to decide which of the two data sets given below has a greater IQR without making any calculations. His answer and reasoning are given below. He has made a mistake in his thinking. Explain the error in his thinking and correct his answer.



*"Data Set B has the larger IQR because its 1<sup>st</sup> and 3<sup>rd</sup> quartiles and median are bigger than the quartiles and medians for Data Set A. Also Data Set B has a larger range."*

6. Create a data set with at least 8 data values that match each set of given conditions.
- a. A data set where the IQR is equal to the range.
  - b. Two different data sets that have the same range but a different IQR.
  - c. A data set with a range of 10 and an IQR of 3.

### 4.3c Class Activity: Box Plots



In the previous section we discussed how to analyze the variability of a data set where the best measure of center is the median. To do so we looked at the *maximum value*, *minimum value*, *median value*, *1<sup>st</sup> quartile*, and *3<sup>rd</sup> quartile*. These values also help us to better understand the distribution of a data set. They are often called **the 5-number summary**. Review what each value in the 5-number summary describes.

#1 Maximum Value

#2 Minimum Value

#3 Median

#4 First Quartile (lower quartile)

#5 Third Quartile (upper quartile)

1. The number of cavities that a class of 6<sup>th</sup> graders have had are shown.

<u>Number of cavities</u>		
0	0	1
2	2	2
3	3	3
4	5	5
7	8	8
12	1	3
3	6	9

- a. Find the 5-number summary for this data set.

A **box plot** is a data display that divides a set of data into four parts using the median and quartiles. It can be used to better help us to understand a data distribution. It is a picture that shows how the 5-number summary values are related to each other on a number line.

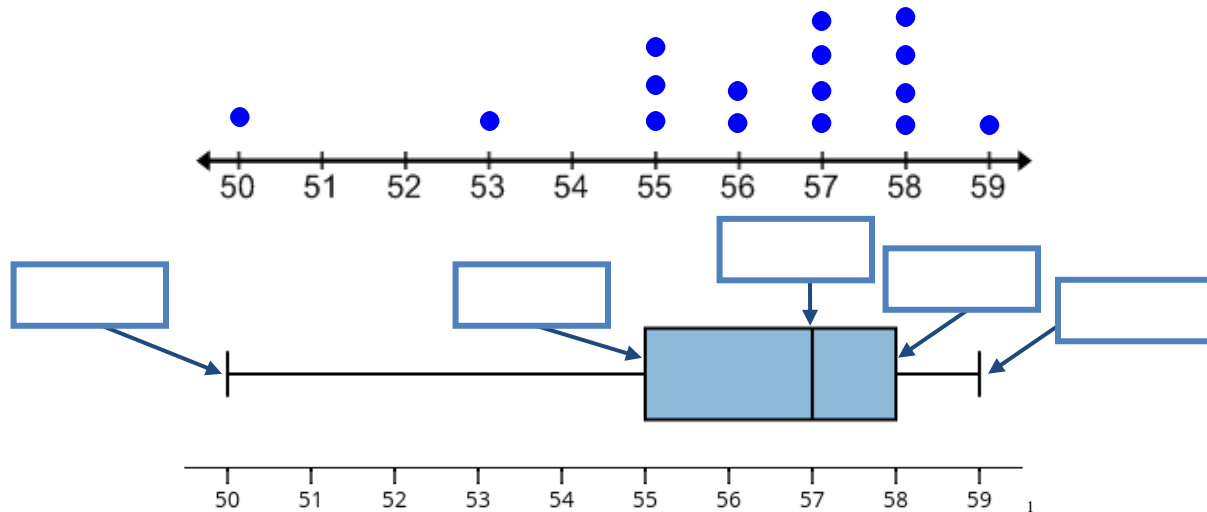
- b. Make a Box Plot of the cavity data on the previous page using the values in its 5-number summary.

2. The data below shows the weight of monkeys at a zoo.

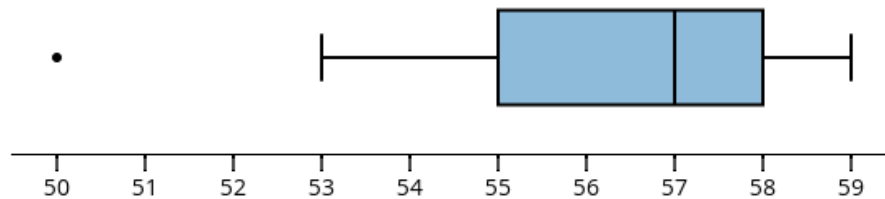
Weight of Monkeys in pounds: 17, 22, 16, 40, 32, 38, 32, 32, 29, 28, 35, 37

- a. Work with a partner to make a box plot of the data. (Hint: begin by finding The 5-Number Summary values of the data.)
- b. Use the box plot to find the IQR of the data. Use the median and IQR to interpret the variability of the data.
- c. What might explain the variability in weight for the monkeys at the zoo?

3. The dot plot and the box plot below represent the same set of data.
- a. Fill in the missing labels with the 5-number summary values on the box plot.



- b. In which display is it easier to identify the median and the quartiles of the data?
- c. In which display is it easier to identify exactly how many data points have a value of 58.
- d. Use the plots to describe the distribution of this data, be sure to discuss the *center*, *shape* and *spread/variability* of the data.
- e. The plot below is another box plot of this data. Compare this plot to the one above, how is it the same? How is it different? Which plot is better?



1

4. Make a box and whisker plot of the following data sets.

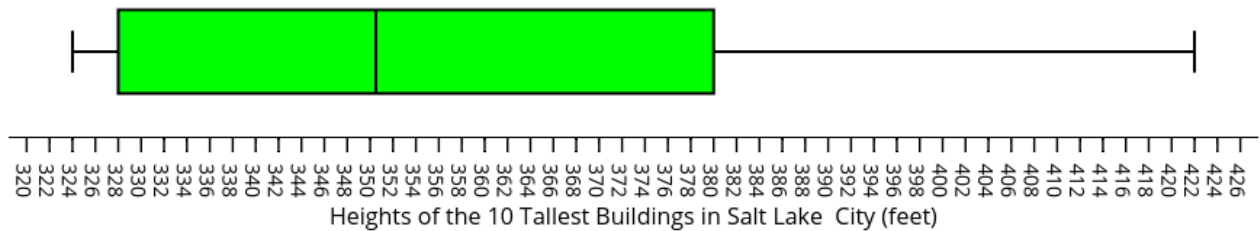
a. Quiz Scores: 18,22, 29, 20, 22, 18, 15, 19, 17, 20 ,18, 19, 11

1

b. Ski Lengths (centimeters): 182, 174, 208, 150, 220, 165, 200, 215, 170, 165, 205

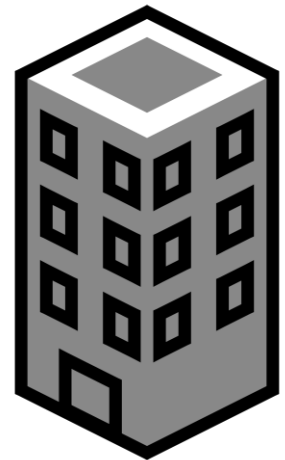
2

5. The box plot shows the height of the tallest buildings in Salt Lake City. Identify if the following statements are true or false. Justify your answer.

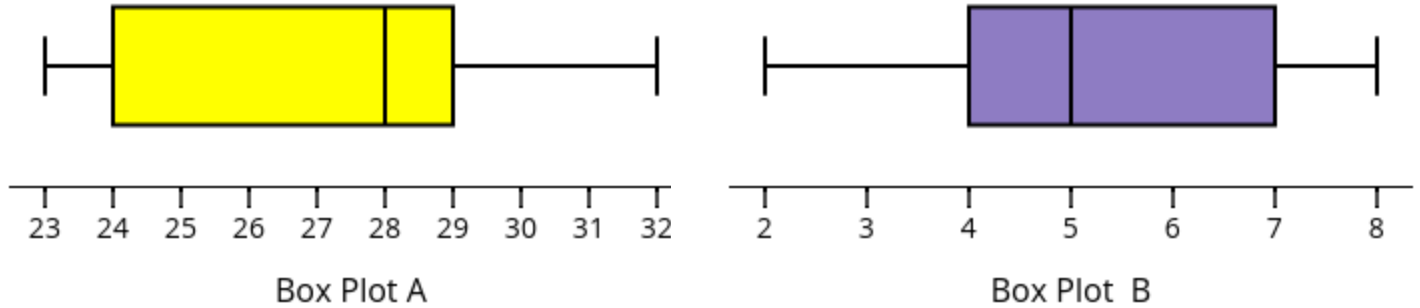


1

- The tallest building in Salt Lake City is 422 feet.
- At least half of the buildings are less than about 350.5 feet tall.
- Half of the buildings are taller than 380 feet tall
- Most of the buildings are greater than 328 feet
- Half of the buildings heights are within 52 feet.
- We know that the value of 422 feet is an outlier because the line segment connecting it is so long.



6. Find the IQR of the two plots given; use the IQRs to determine which data set has less variability around the median.



1

2

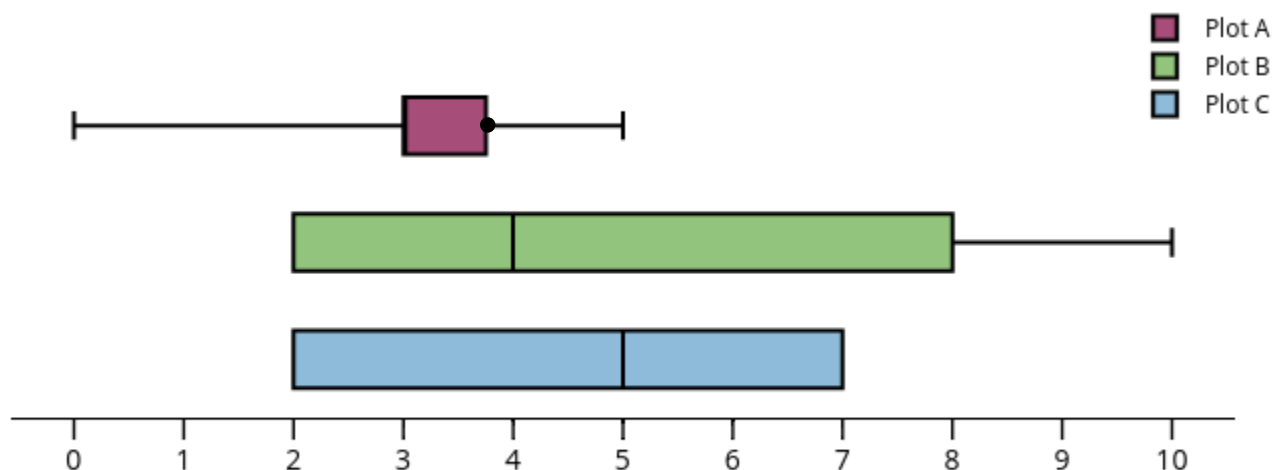


7. Create a box plot for the information given below. Then find the IQR and describe the distribution of the data.

A group of 6<sup>th</sup> graders are interested in the number of hours of TV they watch each week. The most amount of TV watched per week in the group is 22 hours. The least amount of TV watched per week is 0. Half of the group watched less than 10 hours of TV per week. 25% of the group watched more than 15 hours of TV per week and the 1<sup>st</sup> quartile of the data is 8 hours of TV watched per week.



8. Discuss the three box plots below. What makes them different from the plots previously discussed?



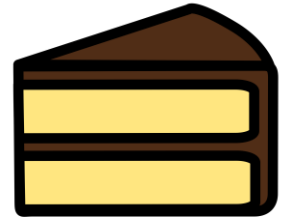
### Spiral Review

- How does the value of the digit of 2 in 2,356 compare to the value of the digit of 2 in 3,256?
  - It is 100 times the value.
  - It is  $\frac{1}{100}$  of the value.
  - It is 10 times the value.
  - It is  $\frac{1}{10}$  of the value.
- How does the value of the digit of 5 in 34.58 compare to the value of the digit of 5 in 85.43?
  - It is 100 times the value.
  - It is  $\frac{1}{100}$  of the value.
  - It is 10 times the value.
  - It is  $\frac{1}{10}$  of the value.
- What decimal number is represented by  $8 \times \frac{1}{10} + 3 \times \frac{1}{100} + 2 \times \frac{1}{1000}$ ?
- Write 34.06 in expanded notation.

### 4.3c Homework: Box Plots

1. The amount of calories per slice in several different types of cake is shown below.

Calories in a Slice of Cake			
Type of Cake	# of Calories	Type of Cake	# of Calories
Carrot	500	Angel Food Cake	70
Chocolate Fudge	450	Lite Coconut Crème	95
Red Velvet	300	White Chocolate Raspberry	340
Lemon Crème	325	Dulce De Leche	350
Devil's Food Cake	350	Hazelnut N Crème	400
German Chocolate	470	Cherry Chocolate Chip	425



- a. Find the 5-Number Summary values for this set of data.

- b. Make a box plot of the data.

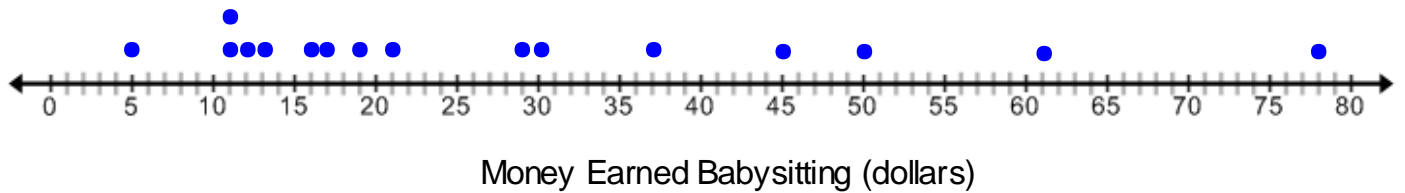
- c. What kind of cake has the most calories per slice?

- d. What percentage of the cakes have at least 350 calories?

- e. Why is the median line not in the center of the box?

- f. If you were on a diet what kind of cake would you choose?
- g. Find and interpret the IQR from the box plot.
- h. Write a few sentences that summarize the distribution of this data. Be sure to discuss the shape, center, and spread.

2. The dot plot shows the amount of money that Tommy has earned babysitting this year.



- a. Make a box plot of the data.

1

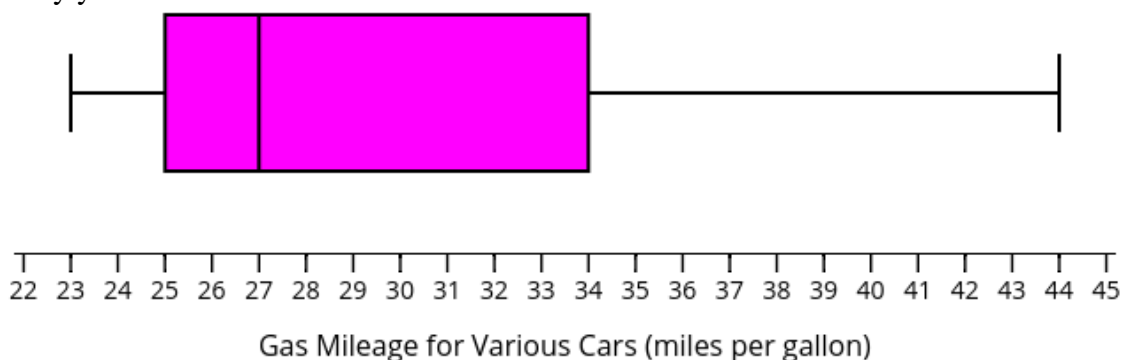
- b. List one way a dot plot is a better display of data than a box plot. Then list one way a box plot is a better display of data than a dot plot.
- c. Write a few sentences describing the distribution of this data. Be sure to discuss the shape, center, and spread.

3. Make a box plot for each set of data below.

a. Hours spent practicing the piano: 5, 8, 5, 10, 5, 8, 7, 9, 10, 3, 0, 8, 7, 6

b. Heart rate in beats per minute: 72, 65, 45, 110, 45, 51, 92, 57, 63, 70, 84

4. The box plot shows the gas mileage for various cars. Identify if the following statements are true or false. Justify your answer.



i. The car with the best gas mileage gets 44 miles per gallon (mpg).

ii. It can be determined with certainty from this plot that at least one car gets 27 mpg.

iii. Half of the cars get less than 34 mpg.

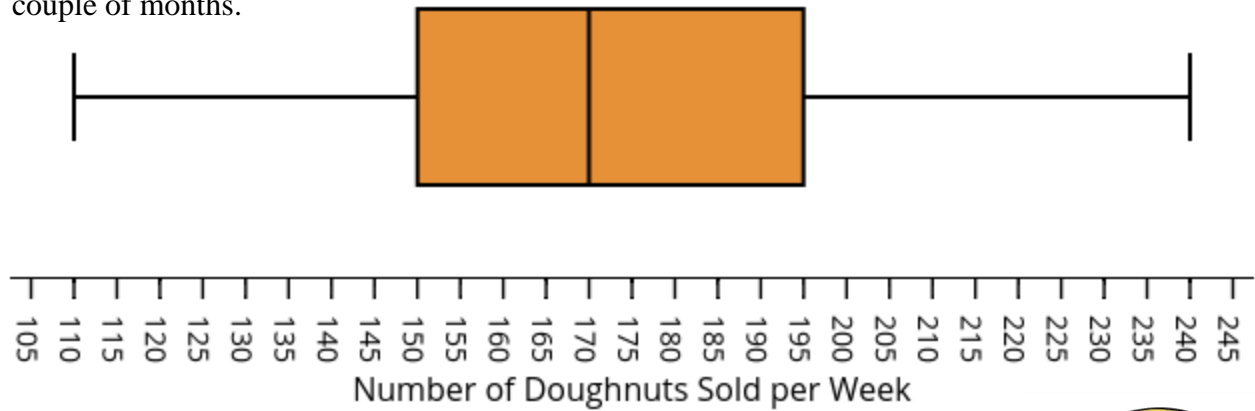
iv. Half of the cars get between 25 and 34 mpg.

v. More data values fall within the top section of the box than the bottom section of the box because the median line is further to the left.

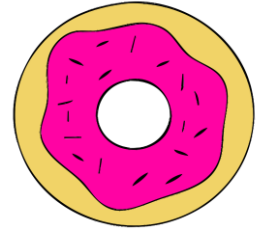
vi. A car that gets 42 mpg is comparable to the gas mileage of most cars.



4. Cindy has made a box plot of the number of doughnuts she has sold each week in her bakery for the last couple of months.



- Make a statement about the data that discusses the center.
- Make a statement about the data that discusses the shape.
- Make a statement about the data that discusses the spread or variability.
- How can Cindy use this information in the future?



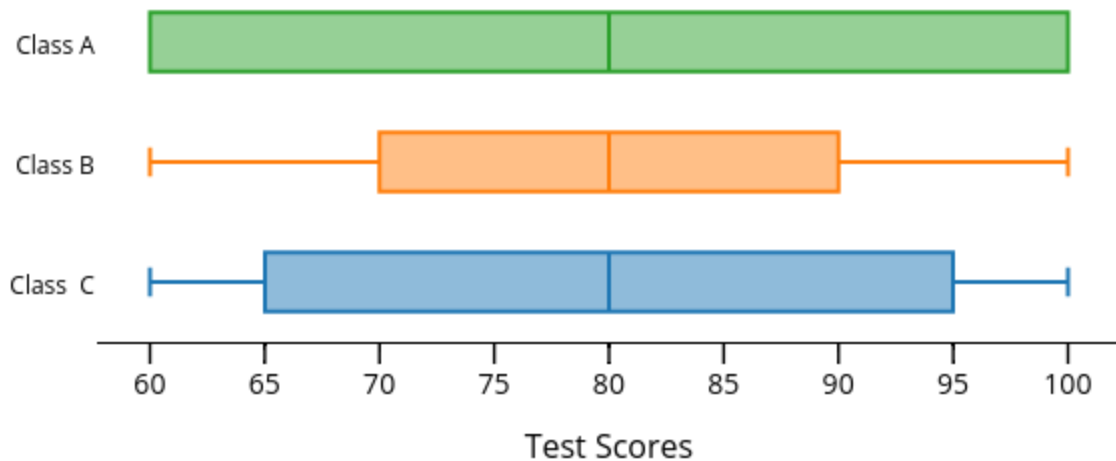
5. Suppose you know the following about a data set: the maximum value is 57, the 3<sup>rd</sup> quartile is 40, the IQR is 15, half of the data is below 27 and the minimum value is 14.
- Make a box plot of the data.

- Are there more data values above or below the median?
- Think of a context for which these values make sense.

### 4.3d Class Activity: Analyzing Box Plots



1. The three box plots below represent the test scores for three different classes. Examine each plot and then discuss the questions that follow.



1

- a. What is the same about these box plots and what is different?
- b. Find the IQR for each plot and use it to compare the variability of each set of class scores.
- c. Make an argument for each class that supports the claim that this class performed the best of the test.

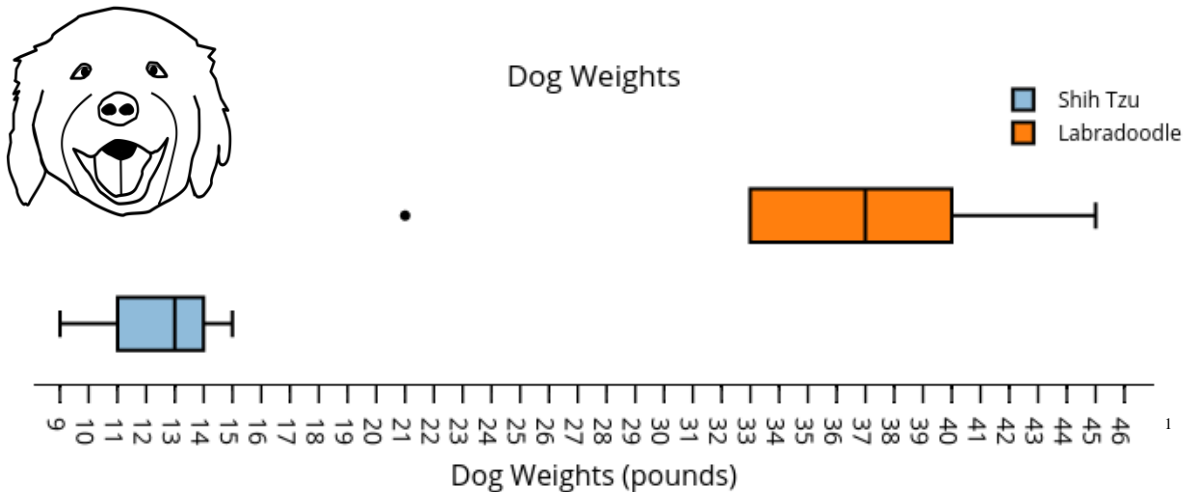
2. Sierra has a small landscaping business. Her monthly profits for the last 12 months are shown in the table.

Month	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
Profit \$	\$1500	\$1500	\$3000	\$4000	\$5500	\$7000	\$7500	\$7000	\$6500	\$5000	\$3000	\$2000

a. Make a box plot of the data.

- b. Write at least 3 bullet points that describe Sierra's profit for her landscaping business. Be sure to discuss the shape, center, and spread of the data

3. The weight, in pounds, of two different breeds of dogs is shown in the box plots.



Write at least 3 bullet points that describe how the weights of the Shih Tzu compare to the weights of Labradoodles?

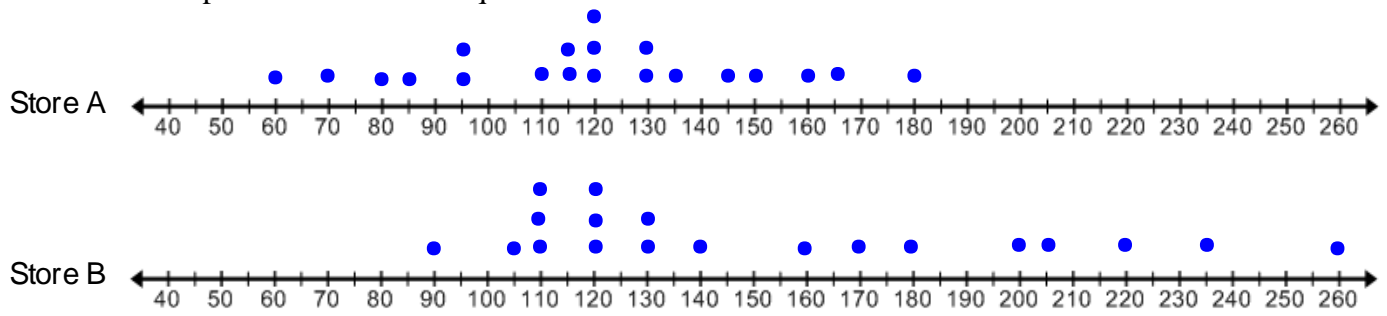
4. The box plots show the height of all the trees in two different city parks.



- How tall is the smallest tree at Centennial Park?
- How tall is the tallest tree in at Valley View Park?
- Which park has trees that are more varied in height? Justify your answer.
- How does the median height of trees in Centennial Park compare to the median height of trees in Valley View Park?  
.
- Give an explanation that could account for the different medians and variability between the heights of trees in these parks.



5. Use the dot plots to answer each question.



a. What does each data point represent?

Price of Guitars (dollars)

b. How many guitars does Store B have in their store?



c. Based off of the dot plots, which store appears to have more variability in prices of guitars. How would this be reflected in their IQR and box plot?

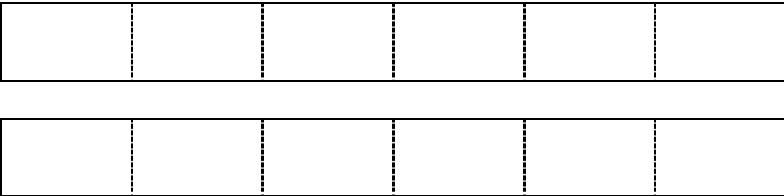
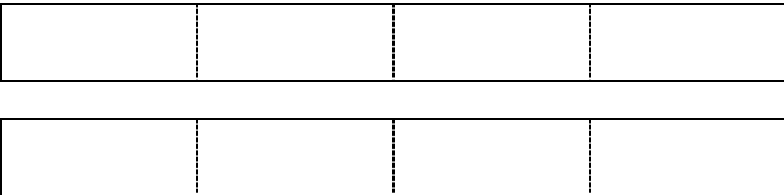
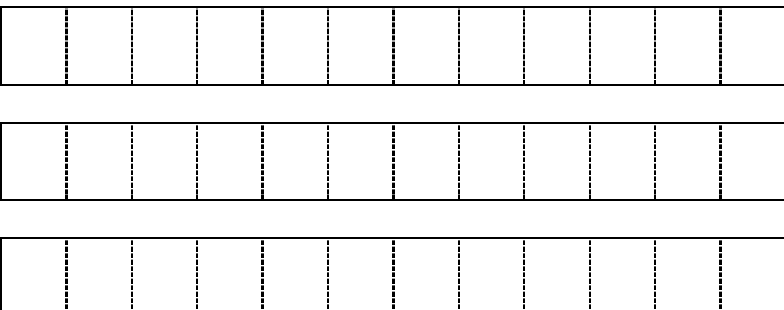
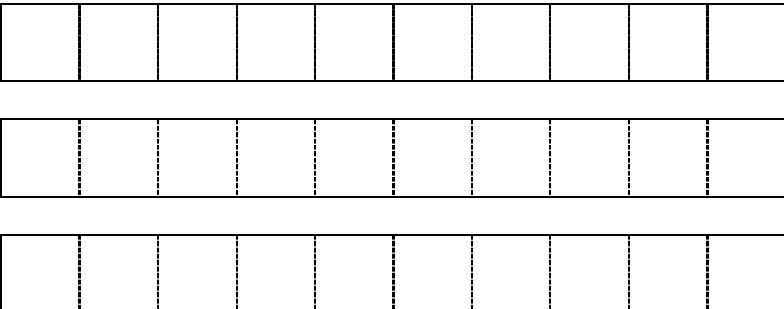
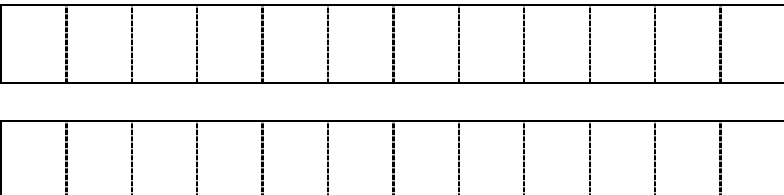
d. Make a box plot for each stores data.

e. Find the Range and IQR for each store and use them to verify which store has more variability in its guitar pricing.

f. Write a few sentences that compare the distribution of data for each store. Be sure to discuss the shape, center, and spread.

# Spiral Review

Directions: Use a visual model to solve each problem.

<p>1. <math>\frac{2}{3} + \frac{1}{2}</math></p>	
<p>2. <math>\frac{3}{4} + \frac{1}{2}</math></p>	
<p>3. <math>1\frac{3}{4} - \frac{1}{3}</math></p>	
<p>4. <math>1\frac{1}{2} + \frac{1}{5}</math></p>	
<p>5. Samantha is cutting fabric. She has a piece that is <math>\frac{7}{12}</math> of a foot long. She cuts off <math>\frac{1}{3}</math> of a foot. How much fabric does she have left on the original piece?</p>	

4.3d Homework: Analyzing Box Plots

1. Create a statistical question of your own that you can answer by surveying people in your class. Then ask at least 15 people in your class your question and record your data below.

a. Write your question here:\_\_\_\_\_

b. Record your data in the table below.

Student	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
Response															

c. What is the attribute that is being measured from your question?

d. What is the unit of measurement for your attribute?

e. Display your data in a box plot, be sure to label your number line and give it a title.

f. What is the median of your data?

g. What is the overall shape of the data?

h. Discuss the variability or spread of the data.

i. Use the distribution of the data in your box plot to answer your question.

### 4.3e Self Assessment: Section 3.3

Consider the following skills/concepts. Rate your comfort level with each skill/concept by checking the box that best describes your progress in mastering each skill/concept. Corresponding sample problems, referenced in brackets, can be found on the following page.

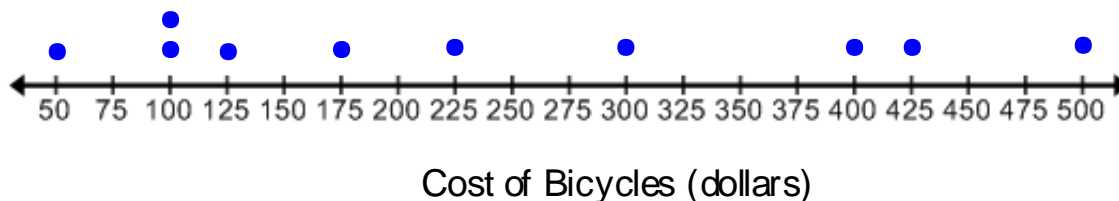
Skill/Concept	Minimal Understanding 1	Partial Understanding 2	Sufficient Mastery 3	Substantial Mastery 4
1. Find and interpret the range for a data set.	I can find the range of a data set from a list or dot plot. I do not know how to interpret the range or find it in a box plot.	I can find the range of a data set that is presented in a list, dot plot, or box plot. I struggle to interpret it from the given context.	I can find and interpret the range of a data set that is presented in a list, dot plot, or box plot. I do not understand why you cannot find the exact range from a histogram.	I can find and interpret the range of a data set that is presented in a list, dot plot, or box plot. I understand and can explain why you cannot find the exact range from a histogram.
2. Find and interpret the mean absolute deviation (MAD) for a data set.	I can sometimes find the MAD but I don't really know why it is useful.	I know how to find the MAD for a data set but cannot interpret its meaning.	I know how to find the MAD for a data set and can interpret its meaning for a given context.	I know how to find the MAD for a data set and can interpret its meaning for a given context. I understand that the MAD is a measure of variability that relates to the mean.
3. Find and interpret values of the 5-number summary for a data set.	I can find some of the 5-number summary values. I don't really know what they mean.	I can find and interpret at least 3 of the 5-number summary values.	I can find and interpret at least 4 of the 5-number summary values.	I can find and interpret all of the 5-number summary values.
4. Find the interquartile range (IQR) for a data set and interpret its meaning.	I can sometimes find the IQR but I don't really know why it is useful.	I know how to find the IQR for a data set but cannot interpret its meaning.	I know how to find the IQR for a data set and can interpret its meaning for a given context.	I know how to find the IQR for a data set and can interpret its meaning for a given context. I understand that the IQR is a measure of variability that relates to the median.

5. Display numerical data in a box plot, use key terms to describe its shape, and interpret the distribution of data.	I can display data in a box plot but often don't know how to label it appropriately. I don't know how to describe the shape or interpret the distribution of data.	I can display data in a box plot that is accurately labeled. I can identify some of the terms used to describe its shape. I struggle to interpret the distribution of data.	I can display data in a box plot that is accurately labeled. I can identify and use terms such as max, min, quartiles, inter-quartile range, symmetrical and outliers to describe the shape and interpret the distribution of data. I cannot determine if the data is skewed.	I can display data in a box plot that is accurately labeled. I can identify and use terms such as max, min, quartiles, inter quartile range, skewed, symmetrical and outliers to describe the shape and interpret the distribution of data.
6. Draw conclusions about two or more data sets by comparing their box plots.	I can draw conclusions about the center when comparing box plots but I don't know how to compare the shape or variability.	I can draw conclusions about the shape, center, and spread when comparing box plots.	I can draw conclusions about the shape, center, and spread when comparing box plots. I take into account the context in which the data is given.	I can draw conclusions about the shape, center, and spread when comparing box plots. I take into account the context in which the data is given. I understand that when comparing box plots I must take into account how they are each scaled.

### Sample Problems for Section 4.3

*Square brackets indicate which concept/skill the problem aligns to.*

1. The dot plot below shows the prices of several bikes as a bike shop. [1][2]



- Find and interpret the range for this data distribution.
- Find and interpret the mean.

- c. Based off of the shape of the data would you say that this distribution has a lot of variability or a little bit of variability? Justify your answer.
- d. Find and interpret the mean absolute deviation (MAD).
- e. At a different bike store the mean price for a bike is also \$240 and the MAD is \$50. Do the prices for bikes at the second store have more or less variability than the prices for bikes at the first store? Justify your answer.
2. A swimming pool maintenance worker is interested in average number of gallons of water that the backyard pools that he services holds. The amount of water that each pool holds is recorded below. [1][3][4]

Amount of water in gallons: 3105, 4075, 1100, 2050, 2325, 4400, 2475, 1500, 3520

- a. Find the values for the 5-Number Summary for this data set.
- Max-
- Min-
- Median-
- Lower Quartile (1<sup>st</sup> quartile)-
- Upper Quartile (3<sup>rd</sup> quartile)-
- b. What percent or fraction of the swimming pools have more than 2475 gallons of water?
- c.  $\frac{1}{4}$  or 25% of the pools have less than how many gallons of water?
- d. What percent of the swimming pools lie between the first quartile and the third quartile?
- e. Find and interpret the range for this data.
- f. Find and the interquartile range (IQR) for this data distribution.

3. The data below shows the heights of all of the roller coasters at an amusement park. [1][4][5]

Height of Roller Coasters (feet)												
203	163	130	260	144	208	150	161	210	222	135	172	230

- a. Make a box plot of the data

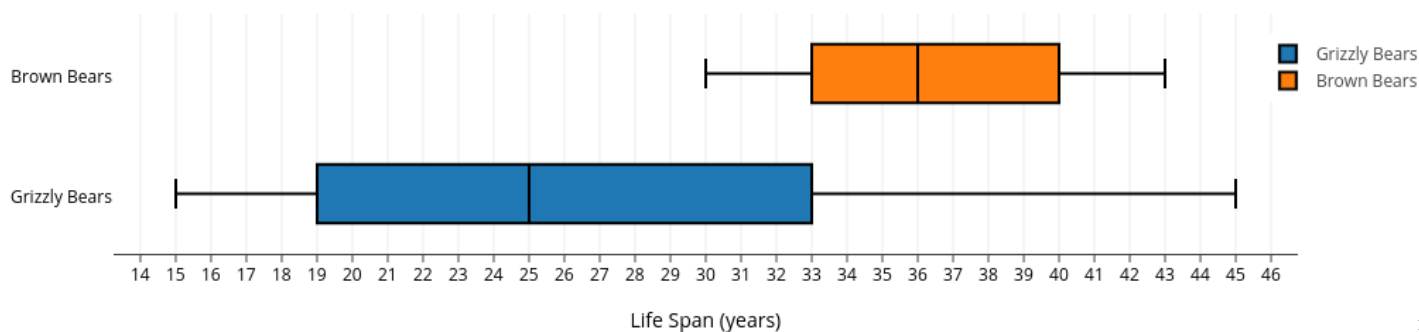
1

- b. What is the median?

- c. Use the box plot to describe the shape of the data distribution.

- d. Find and interpret the range and IQR.

4. The box plots below show the life spans of two different species of bears in a nature preserve. Compare and contrast the ages of the bears. Be sure to discuss shape, center, and variability. [4][6]



1

## Section 4.4: Interpret and Draw Conclusions about Data

### Section Overview:

This section acts as a culmination of the skills of statistical analysis that students have learned throughout the chapter. Students begin to learn how to choose which statistical models and measures will be useful in interpreting data and answering questions. Depending on what is desired, they must choose whether a dot plot, histogram, or box plot best represents the data. They also learn that different information can be obtained from different displays. Understanding when a measure of center or a measure of variability is useful in answering a question is also important. Finally, students are required to develop a deeper understanding of these statistical measures as they compare, interpret, and draw conclusions about several data sets in a variety of contexts, and in turn, collect and analyze data of their own in order to answer a statistical question.

### Primary Concepts and Skills to Master in this Section:

*By the end of this section, students should be able to:*

1. Determine whether a measure of center or a measure of variability should be used to answer a statistical question.
2. Interpret, compare, and draw conclusions about data distributions by using statistical models and appropriate measures of shape, center and spread.

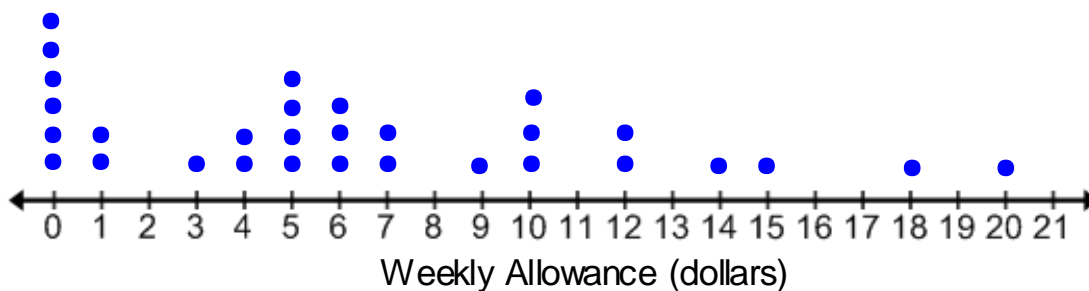
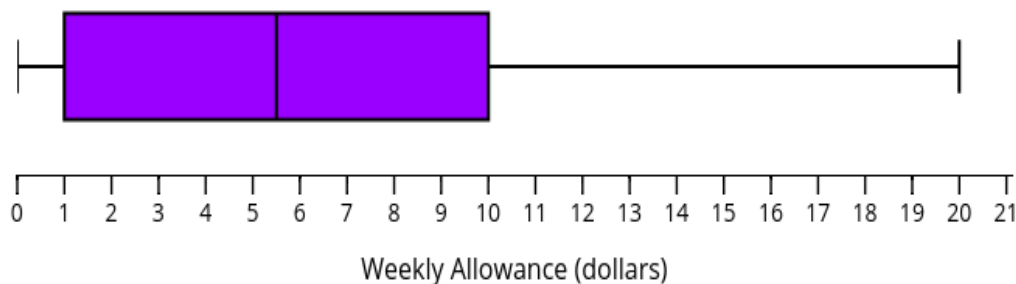
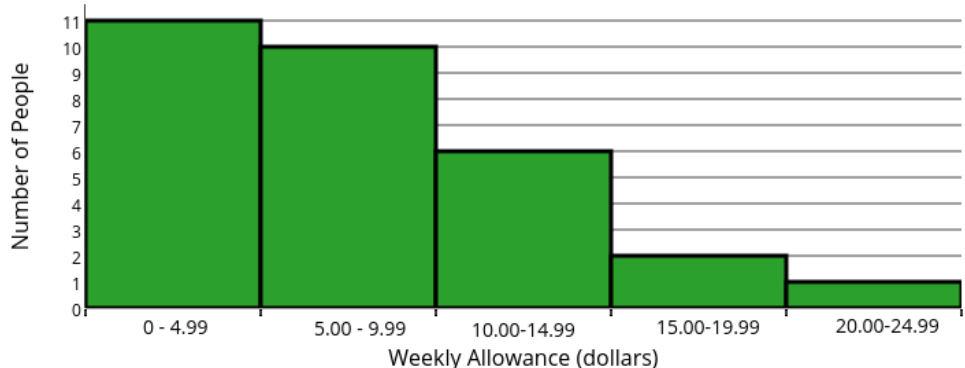


#### 4.4a Class Activity: Make, Interpret and Draw Conclusions About Data Part 1



Donovan, Marcus, and Luther have collected data together on the amount of allowance other students in their class get per week. Each of them displayed the data in a different way. Use their data displays to answer the questions that follow.

Weekly Allowance



1. What is the shape of the data? Which data display did you use to determine this answer? Explain why you chose this display.

2. What is the minimum? Which data display did you use and why?
3. How many people were surveyed? Which data display did you use and why?
4. What is the median of the data? Which data display did you use and why?
5. Describe the variability of the data set. Which data display did you use and why?
6. What is the mean? Which data display did you use and why?

For each of the five statistical questions below (Questions A through E), decide if you would answer the question by considering center or considering variability in the data distribution.

7. A car repair shop keeps track of the number of days it takes to service a car at their shop.

Question A: For cars sent to this shop, what is a typical number of days that it takes for a car to be repaired?

8. Two different cities record their average rainfall for each month during the year.

Question B: On average, which cities receive more rainfall per year?

Question C: Which city has monthly rainfalls that are more consistent (more similar to one another) from month to month?

9. Bags of apples are often measured by weight rather than the number of apples each bag contains. Suppose that you own a grocery store and you carry two different types of apples in your store, Red Delicious and Granny Smiths. For the past year, for each brand of apple, you have kept track of the number of apples contained in each bag delivered to you.

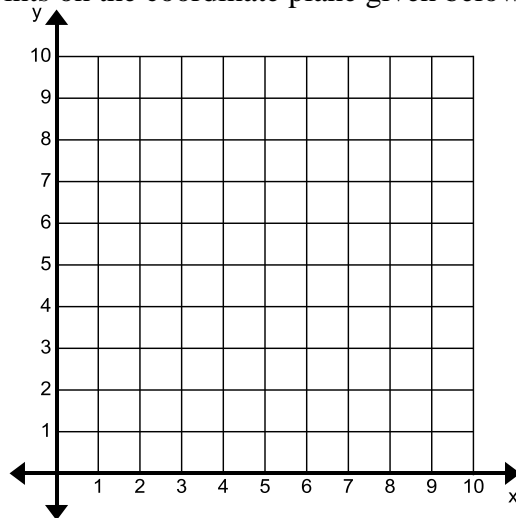
Question D: If you wanted to ensure that your customers purchased bags of apples that each have the same number of apples in them, should you buy Red Delicious apples or Granny Smith apples in the future?

Question E: If you wanted to ensure that your customers purchased bags of apples that contain the greatest number of apples, should you buy Red Delicious apples or Granny Smith apple in the future?

### Spiral Review

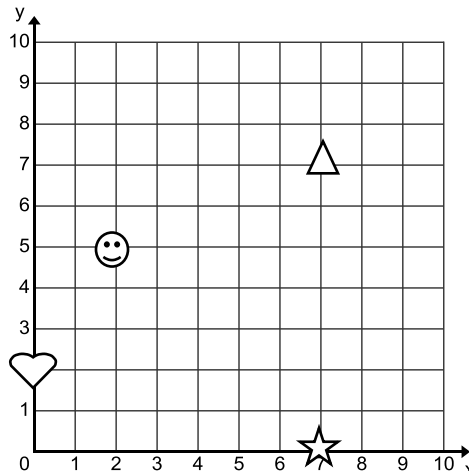
1. Graph and label each set of points on the coordinate plane given below.

- a. A (1,4)
- b. B (0,5)
- c. C (6,6)
- d. D (4,1)



2. Name the ordered pair for each shape shown on the coordinate plane below.

- a. Star
- b. Triangle
- c. Smiley
- d. Heart



#### 4.4a Homework: Make, Interpret, and Draw Conclusions About Data Part 1

1. Below are the ages of 20 people that sing in a choir together.

16, 21, 24, 29, 30, 30, 32, 32, 32, 34, 34, 35, 39, 42, 42, 43, 45, 47, 52, 61

- a. Make an appropriate graph to summarize these ages.

1

2

- b. Describe the distribution of ages for people in this choir. Be sure to describe shape, center and variability.
  
  
  
  
  
  
  
  
  
  
- c. What is the typical age for people in this choir? State which kind of numerical measure you used to answer this question and explain your choice.
  
  
  
  
  
  
  
  
  
  
- c. Are most of the people in the choir around the same age? State which kind of numerical measure you used to answer this question and explain your choice.

#### 4.4b Class Activity: Make, Interpret, and Draw Conclusions About Data Part 2

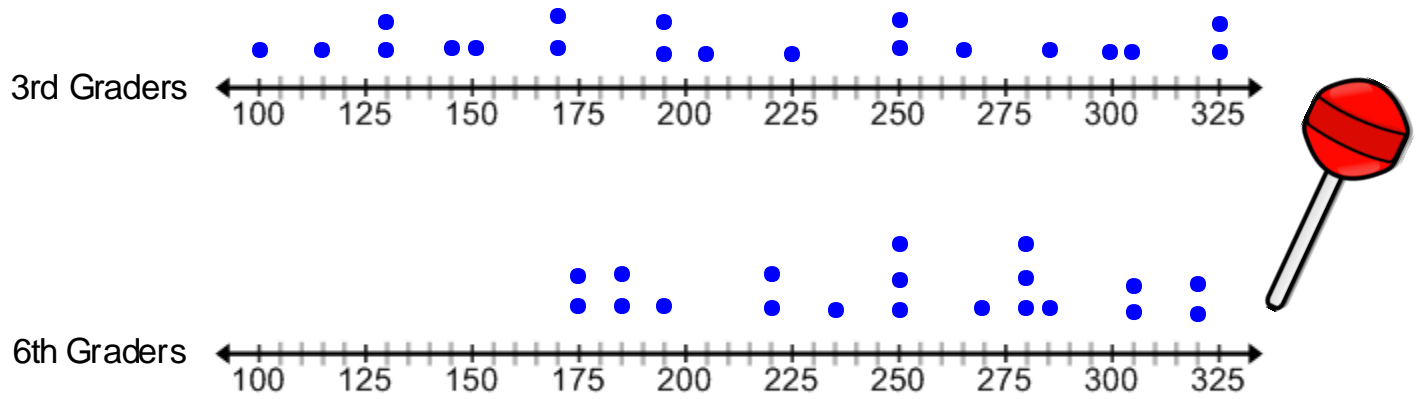


1. Choose the best term from the box to complete each sentence.

Mean	Median	Statistical Question	Mean Absolute Deviation (MAD)	Quartiles	Interquartile Range (IQR)
------	--------	----------------------	-------------------------------	-----------	---------------------------

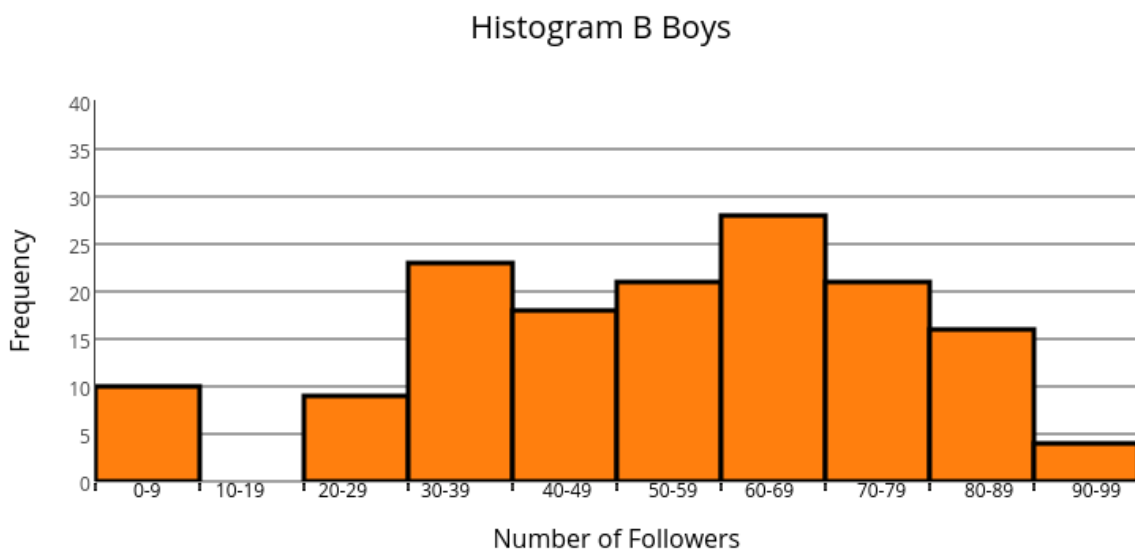
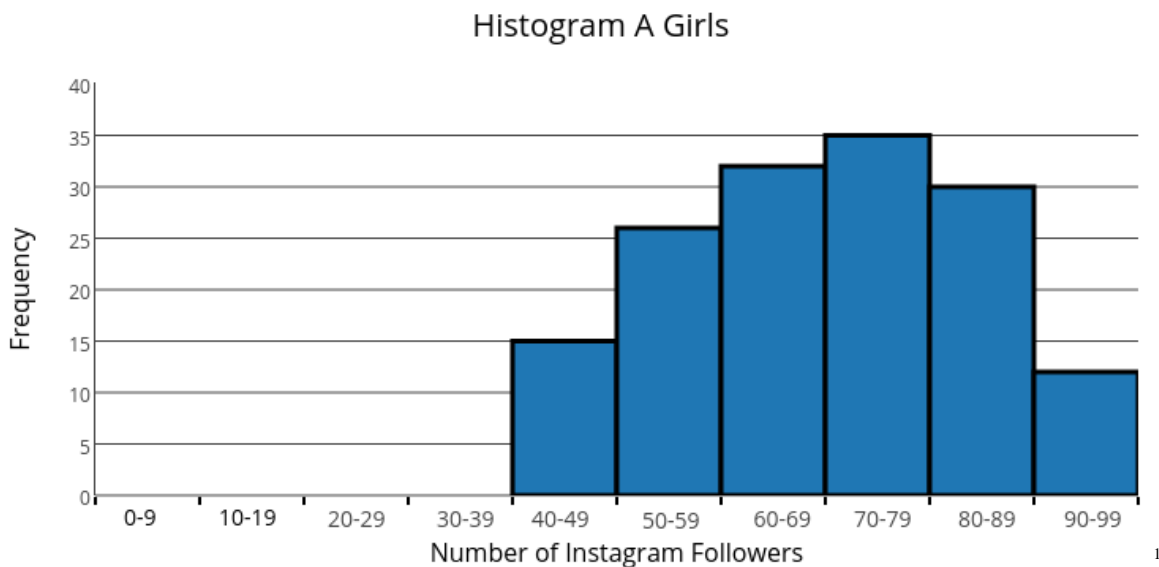
- a. The \_\_\_\_\_ are the values of the points that split the data into four equal parts.
  - b. The \_\_\_\_\_ is a measure of variation. It is computed by finding the *mean* of the absolute deviations in the data set.
  - c. The \_\_\_\_\_ is found by ordering the data from least to greatest and then finding the middle number. If there are two numbers in the middle, then the median is the mean of those two middle numbers.
  - d. A \_\_\_\_\_ generates a variety of answers rather than a single answer.
2. Draw a dot plot with at least 10 data points to match the given conditions.
- a. A dot plot that is symmetrical and has a mean of 6.
  - b. A dot plot that is skewed left and has an outlier.
  - c. A dot plot that has two peaks and a range of 20.

3. The dot plot below shows the number of licks it takes for 20 third graders and 20 sixth graders to get to the center of the Tootsie Pop sucker.



- Describe the data distribution for the 3<sup>rd</sup> graders. Be sure to discuss the shape, center, and spread.
- In general does it take a 3<sup>rd</sup> grader or 6<sup>th</sup> grader more licks to reach the center of a tootsie pop? Justify your answer with statistical measures.
- Which group of students is more consistent in the number of licks that it takes them to get to the center of a tootsie pop? Justify your answer with statistical measures.

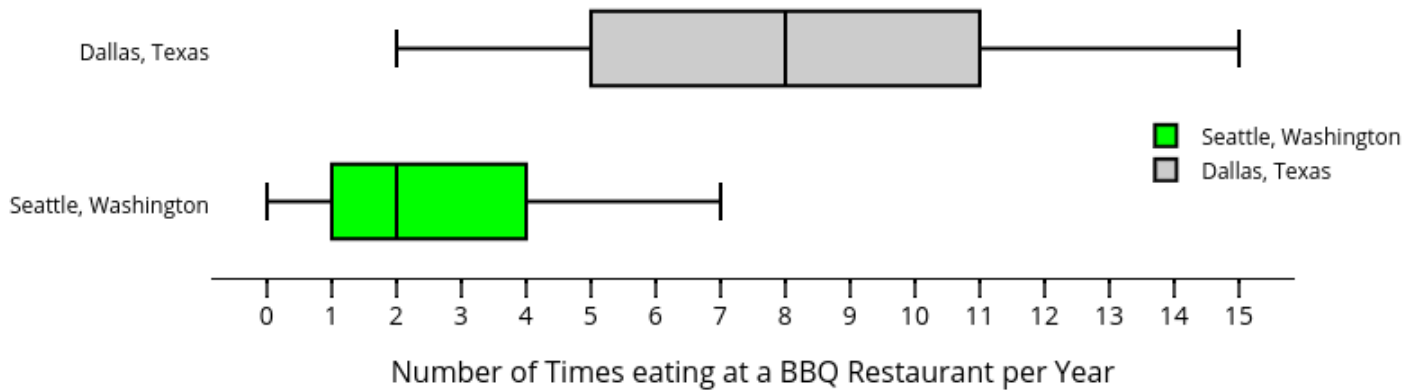
2. Histogram A shows the number of Instagram followers that 150 teenage girls have. Histogram B shows the number of Instagram followers that 150 teenage boys have.



- a. Describe the data distribution for the girls. Be sure to discuss the shape, center, and spread.
- b. On average do girls or boys have more Instagram followers? Is the data for girls and the data for boys centered about the same place? If not, which one has the greater center?
- c. Which group of people has more consistency in the number of Instagram followers that they have?



3. The box plots below consist of data on the number of times that 50 people who live in two major cities eat at a BBQ restaurant each year.

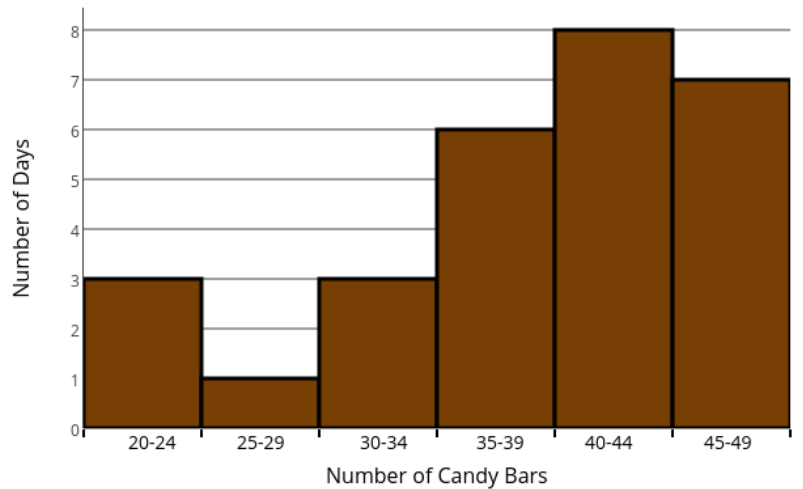


- a. Describe the data distribution for the Seattle Washington. Be sure to discuss the shape, center, and spread.<sup>1</sup>
- b. Do more people typically eat at BBQ restaurants that live in Dallas, Texas or that live in Seattle, Washington? Justify your answer with statistical measures. Is the data for Seattle and the data for Dallas centered in about the same place? If not, which one has the greater center?
- c. Which city has more consistency in the number of people that eat at BBQ restaurants?
- d. Why might this information be useful?

<sup>1</sup> evatipp. "BBQ Restaurants" Plotly Technologies Inc., <https://plot.ly/505/~evatipp/> (Last Accessed: March 2016)

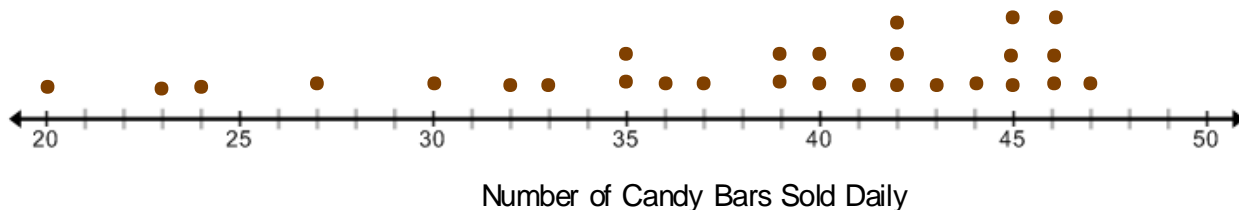
4. Roman thinks that his school needs to get another vending machine in the school cafeteria because there is always a long line to use the machine at lunch time. The principal has told everyone that they cannot get another vending machine unless they can show that on average at least 40 candy bars are sold each day. Roman conducts a survey and recorded his results in the histogram below.

Daily Number of Candy Bars Sold from the Vending Machines



- a. What possible arguments could Roman give to his principal that on average at least 40 candy bars are sold from the vending machine each day?

- b. Roman decides to present his data in a dot plot in addition to the histogram. Use the dot plot to find the exact *mean* and median number of candy bars sold.



- c. In addition to looking at measures of center, Roman wants to consider variability. Find numerical summaries that can help you discuss the variability of the data. Interpret what these numerical summaries mean.
- d. Use the information found in the previous questions to refine your argument about convincing Roman's principal that in general at least 40 candy bars are sold each day.

### Spiral Review

Mark each statement as true or false. If the statement is false, rewrite it so that it is true.

	Statement	True	False
1.	All trapezoids are quadrilaterals.		
2.	All rectangles are squares.		
3.	Kites are never rhombuses.		
4.	Rectangles are always parallelograms.		
5.	All squares are kites.		

#### 4.4b Homework: Make, Interpret, and Draw Conclusions About Data Part 2

1. Choose the best term from the box to complete each sentence

Mean	Histogram	Range	Interquartile Range (IQR)	Dot Plot	Box Plot
------	-----------	-------	---------------------------	----------	----------

- a. A \_\_\_\_\_ is a data display that divides a set of data into four parts using the median and quartiles.
- b. The \_\_\_\_\_ is the difference of the 3<sup>rd</sup> quartile and the 1<sup>st</sup> quartile.
- c. The \_\_\_\_\_ is a measure of center for a data set found by taking the sum of the data divided by the number of data values.
- d. A \_\_\_\_\_ use a number line to show the number of times each value in a data set occurs.

2. Draw a box plot to match the given conditions.

- a. A box that is symmetrical and has a median of 10.

1

- b. A box plot that is skewed left and has an outlier at 5.

2

- c. A box plot that has and IQR of 15 and a range of 30.

3

3. Create a statistical question of your own that you can answer by surveying people in your class. Then ask at least 15 people in your class your question and record your data below.

a. Write your question here:\_\_\_\_\_

b. Record your data in the table below.

Student	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
Response															

c. What is the attribute that is being measured from your question?

d. What is the unit of measurement for your attribute?

e. Summarize your data with a graphical display (histogram, dot plot, or box plot).

f. What is the overall shape of the data?

g. What is the center of your data?

h. Discuss the variability or spread of the data.

i. Use the numerical summaries and graphical display of your data to answer your statistical question.

#### 4.4c Self Assessment: Section 3.4

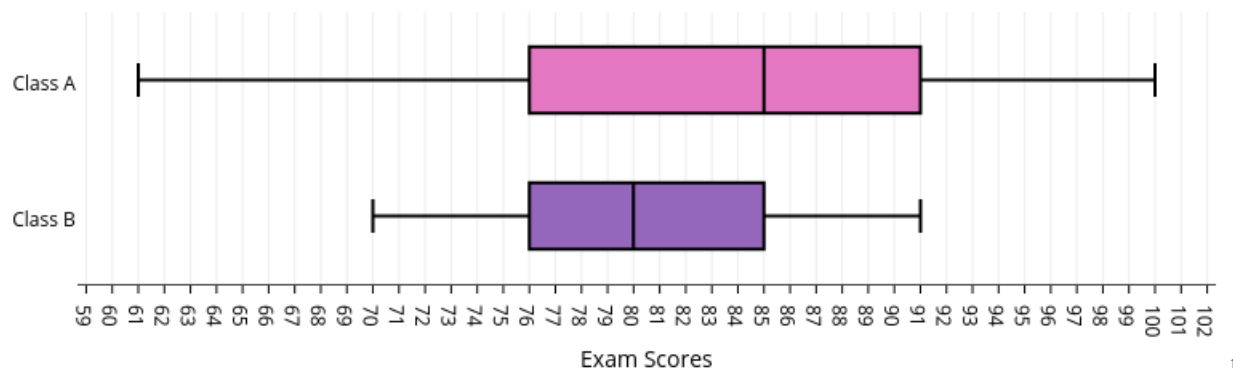
Consider the following skills/concepts. Rate your comfort level with each skill/concept by checking the box that best describes your progress in mastering each skill/concept. Corresponding sample problems, referenced in brackets, can be found on the following page.

<b>Skill/Concept</b>	<b>Minimal Understanding 1</b>	<b>Partial Understanding 2</b>	<b>Sufficient Mastery 3</b>	<b>Substantial Mastery 4</b>
1. Determine whether a measure of center or a measure of variability should be used to answer a statistical question.	I can determine whether a measure of center should be used to answer a statistical question.	I can determine whether a measure of variability should be used to answer a statistical question.	I can determine whether a measure of center or measure of variability should be used to answer a statistical question.	I can determine whether a measure of center or measure of variability should be used to answer a statistical question and in turn use this measure to answer the statistical question.
2. Interpret, compare, and draw conclusions about data distributions by using statistical models and appropriate measures of shape, center and spread.	I can find some measures of shape, center, and spread but I don't really know how to use them to interpret data.	I know how to make some conclusions about a data distribution using statistical models but I struggle to identify and find all the appropriate measures of shape, center, and spread.	I know how to interpret and draw conclusions about a single data distribution by using statistical models and appropriate measures of shape, center, and spread.	I know how to interpret, compare, and draw conclusions about data distributions by using statistical models and appropriate measures of shape, center, and spread.

## Sample Problems for Section 4.4

Square brackets indicate which concept/skill the problem aligns to.

1. The data below shows the recent exam scores for two different English classes. For each of the three statistical questions below (Questions A, B, and C), decide if you would answer the question by considering center or considering variability in the data distribution. Then use this measure to answer the statistical question. [1]



Question A: What is the typical test score for all of the students on this exam?

Question B: Which class has the highest median score for this exam?

Question C: Which class has scores that are more consistent (more similar) with one another?

2. A group of 6<sup>th</sup> graders were each given one minute to make a paper clip chain as long as possible. The number of paper clips on each person's chain is shown below. [2]

11, 13, 16, 16, 17, 18, 21, 21, 22, 22, 22, 23, 24, 24, 26, 26, 28, 28, 29, 31, 32, 36

- a. Choose an appropriate graphical display for your data. Represent the data in this display and describe the shape of the data.

- b. Choose and find an appropriate measure of center for the data. Explain why you chose this measure of center to represent the data and interpret its meaning given the context.
- c. Choose and find an appropriate measure of variability. Explain why you chose this measure of variability to represent the data and interpret its meaning given the context.



This book is shared online by Free Kids Books at <https://www.freekidsbooks.org> in terms of the creative commons license provided by the publisher or author.

Want to find more books like this?



<https://www.freekidsbooks.org>

Simply great free books -

Preschool, early grades, picture books, learning to read,  
early chapter books, middle grade, young adult,

Pratham, Book Dash, Mustardseed, Open Equal Free, and many more!

***Always Free – Always will be!***

**Legal Note:** This book is in CREATIVE COMMONS - Awesome!! That means you can share, reuse it, and in some cases republish it, but only in accordance with the terms of the applicable license (not all CCs are equal!), attribution must be provided, and any resulting work must be released in the same manner.

Please reach out and contact us if you want more information:

<https://www.freekidsbooks.org/about> Image Attribution: Annika Brandow, from You! Yes You! CC-BY-SA. This page is added for identification.