

**Accelerated Algebra 1/Geometry A**  
**Unit 6: Describing Data**

**Enduring Understandings:**

- Data are gathered, displayed, summarized, examined, and interpreted to discover patterns and deviations from patterns.
- Which statistics to compare, which plots to use, and what the results of a comparison might mean, depend on the question to be investigated and the real-life actions to be taken.
- Understand and be able to use the context of the data to explain why its distribution takes on a particular shape (e.g. are there real-life limits to the values of the data that force skewness?)
- When making statistical models, technology is valuable for varying assumptions, exploring consequences and comparing predictions with data.
- Causation implies correlation yet correlation does not imply causation.

**Essential Questions:**

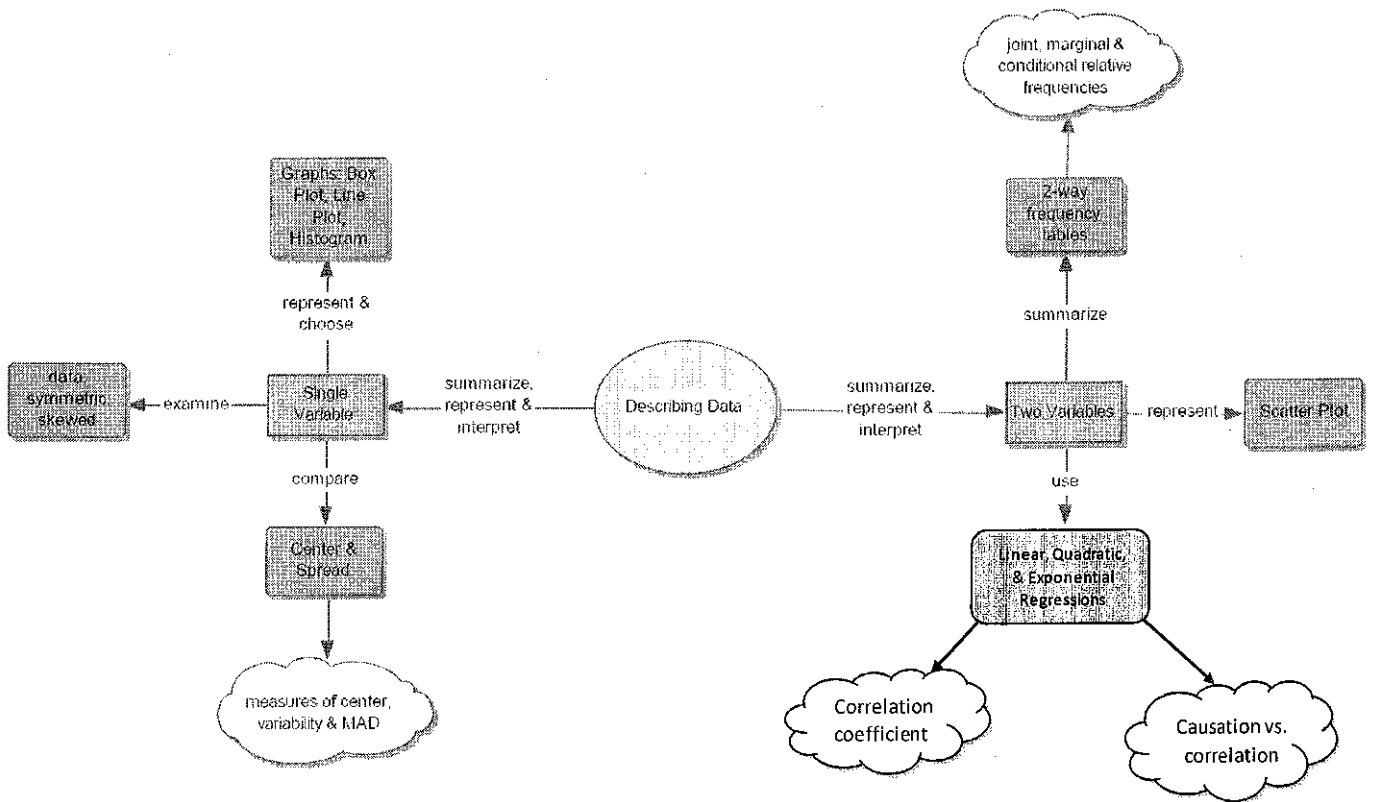
How can I represent, interpret and compare sets of data?

Date	Topic	Assignment
Tuesday January 29	-Dot Plots, Histograms, Box Plots - Measures of Central Tendency Notes	Pages 1-7
Wednesday January 30	-Continue Measures of Central Tendency -Interpret Shapes, Centers and Spread of Data	Pages 8-12
Thursday January 31	Two Way Frequency Tables	Pages 13-20
Friday February 1	<b>QUIZ/Review</b>	Review on own.....catch up on homework
Monday February 4	-Interpret Correlation Coefficient Go over quiz Power point Use calculator to find correlation given data from class (height vs show size) - Correlation Vs. Causation	Pages 21-29
Tuesday February 5	-Lines of best fit Group work: pages 26-27 - Line of best fit (exponential and linear)	Pages 30-34
Wednesday February 6	Residuals	Pages 35-37

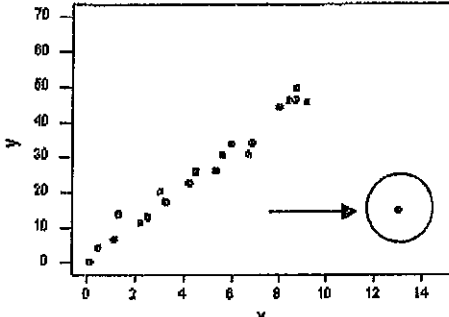
Thursday February 7	Residuals/Review for test	Pages 38-42
Friday February 8	<b>Unit 6 TEST</b>	EOC Review- Unit 1 (odds)
Monday February 11	Unit 6 Touchstone test	EOC Review-Unit 2 (odds)
Tuesday February 12	EOC Review - Unit 3 (odds)	EOC Review – Unit 4 (odds)
Wednesday February 13	EOC Review – Unit 5 (odds)	EOC - Mixed Review (all)
Thursday February 14	<b>EOC</b>	Please bring in board games to do in class tomorrow!! Also, bring your favorite snack to share!
Friday February 15	Game day!	Rest up!!!



Mid-Winter Break!!!!!!!!!!!!!!!!!!!!!! Yay!!!!





<b>Interquartile Range</b>		<b>Subtract</b> <b>Third Quartile (<math>Q_3</math>) - First Quartile (<math>Q_1</math>) = IQR</b>
<b>Outlier</b>		
<b>Mean</b>		$5 + 4 + 2 + 6 + 3 = 20$ $\frac{20}{5} = 4$ <b>The Mean is 4.</b>
<b>Mean Absolute Deviation (MAD)</b>		<b>Steps:</b> <ol style="list-style-type: none"> <li>1. Find the Mean</li> <li>2. Calculate the absolute value of the difference between each data value and the mean</li> <li>3. Determine the average of the differences in step 2. This average is the mean absolute deviation</li> </ol>
<b>Measures of Center</b>		Find the Mean and Median for the following data. <b>Hint:</b> (Must order the numbers first before finding the Median)  2 1 5 4 3  <b>Mean:</b> $\frac{15}{5} = 3$ <b>Median = 3</b>
<b>Measures of spread</b>		<b>Examples of Measures of Spread:</b> <ol style="list-style-type: none"> <li>1. Range</li> <li>2. Interquartile Range (IQR)</li> <li>3. Mean Absolute Deviation -MAD</li> </ol>

Name: \_\_\_\_\_ Date: \_\_\_\_\_

### Graphical Displays for Data

**Example 1:** A pharmacy records the number of customers each hour that the pharmacy is open. The staff is using the information to determine how many people need to be working at the pharmacy at each time of the day. The number of customers is in the table below. Use the table to create a histogram to help the pharmacy staff understand how many customers are in the pharmacy at each time of day.

Time Frame	Number of customers
8:00 A.M. – 9:00 A.M.	2
9:00 A.M. – 10:00 A.M.	0
10:00 A.M. – 11:00 A.M.	8
11:00 A.M. – 12:00 P.M.	14
12:00 P.M. – 1:00 P.M.	23
1:00 P.M. – 2:00 P.M.	12
2:00 P.M. – 3:00 P.M.	7
3:00 P.M. – 4:00 P.M.	3
4:00 P.M. – 5:00 P.M.	5

**Example 2:** Anna and Ethan watch 10 thirty-minute shows during the month of June. They record the number of food commercials that air during each show in the table below. Create a dot plot to display the number of food commercials that aired during the 10 shows.

Shows	# of Commercials
A	7
B	7
C	5
D	7
E	4
F	7
G	5
H	9
I	5
J	6

**Example 3:** Ray's scores on his mathematics tests were 70, 85, 78, 90, 84, 82, and 83. Draw a box plot to represent Ray's Data.

Find the IQR.

Are there any outliers?

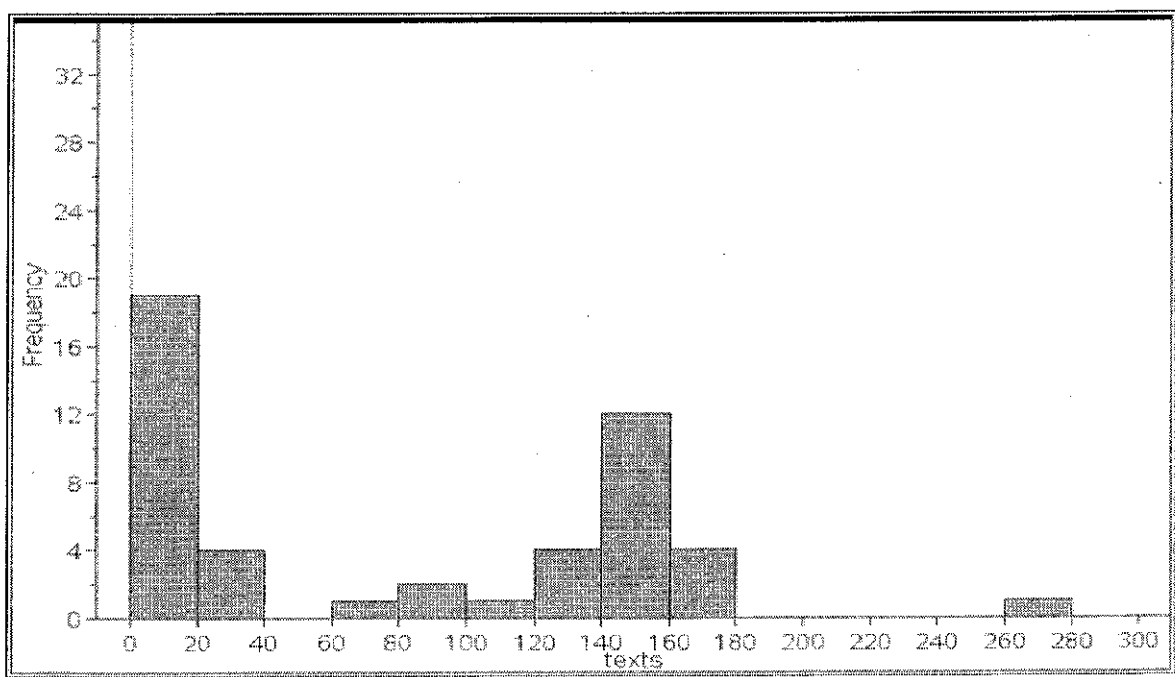


# 8.1 Texting by the Numbers

## A Solidify Understanding Task

Technology changes quickly and yet has a large impact on our lives. Recently, Rachel was busy chatting with her friends via text message when her mom was trying to also have a conversation with her. Afterward, they had a discussion about what is an appropriate amount of texts to send each day. Since they could not agree, they decided to collect data on the number of texts people send on any given day. They each asked 24 of their friends the following question: "What is the average number of texts you SEND each day?" The data and histogram representing all 48 responses are below:

{150, 5.5, 6, 5, 3, 10, 150, 15, 20, 15, 6, 5, 3, 6, 0, 5, 12, 25, 16, 35, 5, 2, 13, 5, 130, 145, 155, 150, 162, 80, 140, 150, 165, 138, 175, 275, 85, 137, 110, 143, 138, 142, 164, 70, 150, 36, 150, 150}



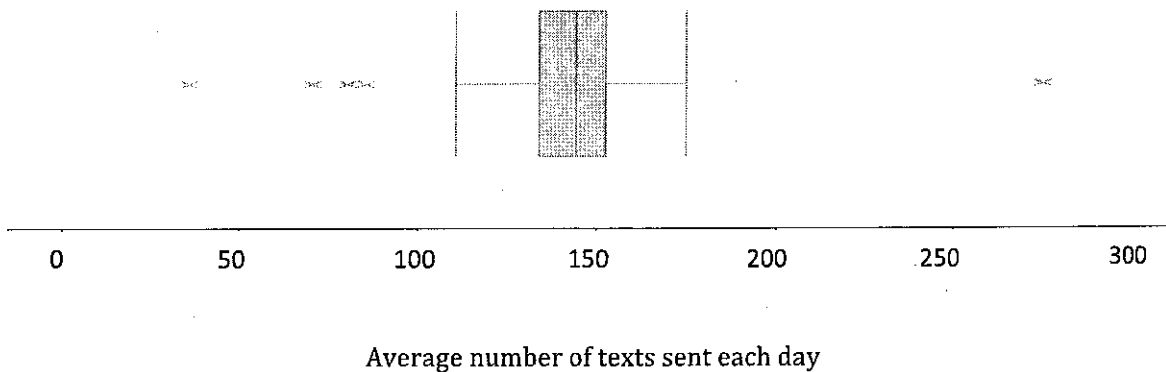
Part I: What information can you conclude based on the histogram above?

Represent the same data by creating a box plot below.

http://www.flickr.com/photos/garryknight/740038165

Describe the pros and cons of each representation (histogram and box plot). In other words, what information does each representation highlight? What information does each representation hide or obscure?

Part II: Prior to talking about the data with her mom, Rachel had created a box plot using her own data she collected and it looked quite different than when they combined their data.



Describe the data Rachel collected from her friends. What does this information tell you?

What do you think is a reasonable number of texts Rachel can send per day?

Rachel wants to continue sending her normal number of texts (average of 100 per day) and her mom would like her to decrease this by half. Present an argument for each side, using mathematics to justify each person's request.

© 2012 Mathematics Vision Project | MVP

In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.



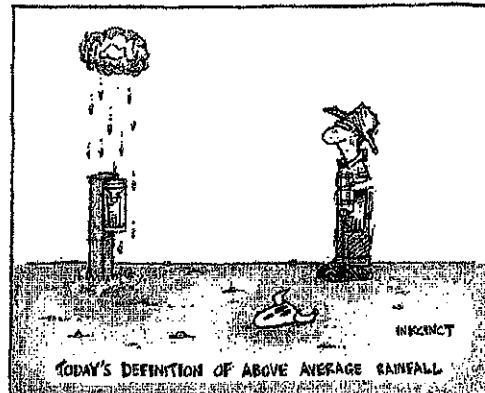
Name: \_\_\_\_\_ Date: \_\_\_\_\_

### Measures of Central Tendency

1. Some people use "average" interchangeably for both mean and median. Consider this statement:

"Just think of how stupid the average person is, and then realize half of them are even stupider?" George Carlin

What type of "average" is George Carlin referring to, mean or median? Is it possible to have more than half of a population above this kind of average?



2. What is the difference between mean and median?
3. Give an example of data when the mean and median might have the same value.

Give an example when the mean and the median do NOT have the same value.

4. Can the following statement be true? Why or why not?  
 "Welcome to Lake Wobegon, where all the women are strong, all the men are good-looking, and all the children are above average." Garrison Keillor
5. Is it possible to have more than half of data values above (or below) the mean?
6. Find the mean, median, and mode for this set of data.  
 5, 11, 16, 8, 4, 7, 15, 6, 11, 7
7. Kara had 85, 83, 92, 88, and 69 on her first five math tests. She knows that she needs an average of 85 to get a B. What score must she get on her last test to get a B?

## 50 | Mean Absolute Deviation

The mean absolute deviation or M.A.D. measures the spread of a set of data just like the interquartile range (IQR). Unlike IQR, however, M.A.D. uses every data point. Because they both use every data point, the mean (center) and M.A.D. (spread) tend to be used together to describe a set of data. These two measures of center and spread are appropriate for symmetric distributions.

### Finding the mean absolute deviation

1. Find the mean of the data
2. Find the difference of every data point from the mean (called deviation)
3. Make every difference positive (absolute value)
4. Find the mean of the absolute differences

**Find the M.A.D. for the following data sets. Round the answers to the nearest tenth.**

1. 68, 70, 72, 73, 74, 75
2. 72, 75, 73, 99, 68, 79, 48, 60, 52, 59
3. 250, 300, 200, 400, 650, 225, 760, 1215
4. 22, 31, 57, 29, 62, 24
5. 1, 1, 3, 3, 6, 6, 5, 5, 10, 12



Name:

## Modeling Data 8.1

Ready, Set, Go!



<http://www.flickr.com/photos/garryknight/740038>

Ready

Topic: Measures of central tendency

**Sam's test scores for the term were 60, 89, 83, 99, 95, and 60.**

1. Suppose that Sam's teacher decided to base the term grade on the mean.
  - a. What grade would Sam receive?
  - b. Do you think this is a fair grade? Explain your reasoning.
2. Suppose that Sam's teacher decided to base the term grade on his median score.
  - a. What grade would Sam receive?
  - b. Do you think this is a fair grade? Explain your reasoning.
3. Suppose that Sam's teacher decided to base the term grade on the mode score.
  - a. What grade would Sam receive?
  - b. Do you think this is a fair grade? Explain your reasoning.
4. Aiden's test scores for the same term were 30, 70, 90, 90, 91, and 99. Which measure of central tendency would Aiden want his teacher to base his grade on? Justify your thinking.
5. Most teachers base grades on the mean. Do you think this is a fair way to assign grades? Why or why not?

© 2012 Mathematics Vision Project | MVP

In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.

Name:

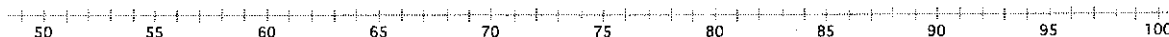
## Modeling Data 8.1

## Set

Topic: Examining data distributions in a box-and-whisker plot

60, 64, 68, 68, 72, 76, 76, 80, 80, 80, 84, 84, 84, 84, 88, 88, 88, 92, 92, 96, 96, 96, 96, 96, 96, 96, 100, 100

6. Make a box-and-whisker plot for the following test scores.



7a. How much of the data is represented by the box?

b. How much is represented by each whisker?

8. What does the graph tell you about student success on the test?

## Go

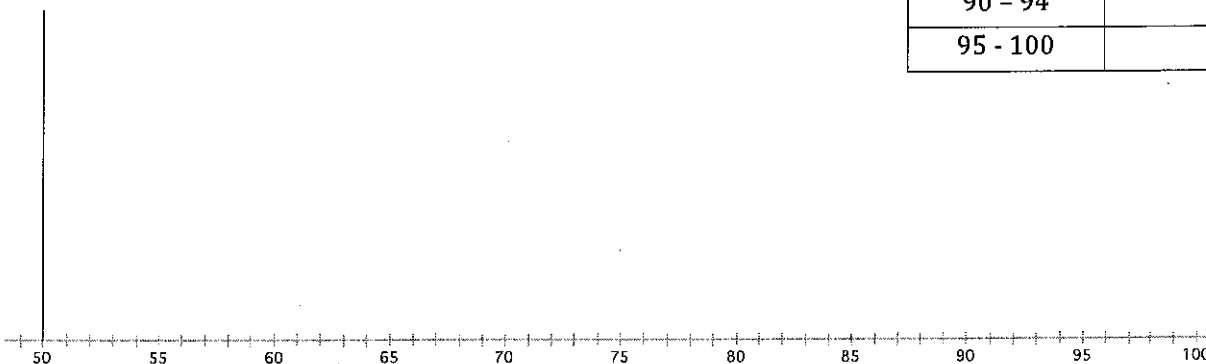
Topic: Drawing histograms.

Use the data from the SET section to answer the following questions

9. Make a frequency table with intervals. Use an interval of 5.

10. Make a histogram of the data using your intervals of 5.

Score	Frequency
60 - 64	
65 - 69	
70 - 74	
75 - 79	
80 - 84	
85 - 89	
90 - 94	
95 - 100	



Need Help? Check out these related videos:

[http://www.khanacademy.org/math/statistics/e/mean\\_median\\_and\\_mode](http://www.khanacademy.org/math/statistics/e/mean_median_and_mode)
<http://www.khanacademy.org/math/algebra/ck12-algebra-1/v/box-and-whisker-plot>
<http://www.khanacademy.org/math/algebra/ck12-algebra-1/v/histograms>

© 2012 Mathematics Vision Project | MVP



In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.

Name: \_\_\_\_\_ Date: \_\_\_\_\_

---

## How to Compare Distributions

---

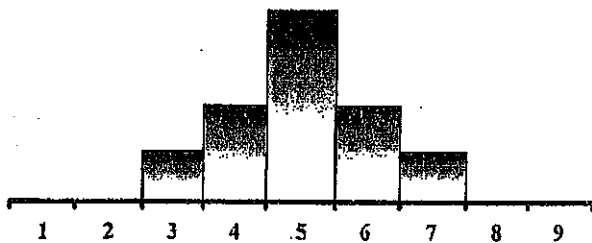
When you compare two or more data sets, focus on four features:

- ★ **Center.** Graphically, the center of a distribution is the point where about half of the observations are on either side.
- ★ **Spread.** The spread of a distribution refers to the variability of the data. If the observations cover a wide range, the spread is larger. If the observations are clustered around a single value, the spread is smaller.
- ★ **Shape.** The shape of a distribution is described by symmetry, skewness, number of peaks, etc.
- ★ **Unusual features.** Unusual features refer to gaps (areas of the distribution where there are no observations) and outliers.

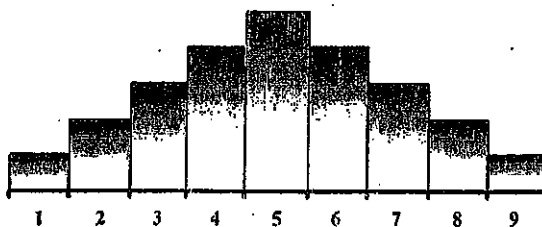
---

### SPREAD

The spread of a distribution refers to the variability of the data. If the data cluster around a single central value, the spread is smaller. The further the observations fall from the center, the greater the spread or variability of the set.



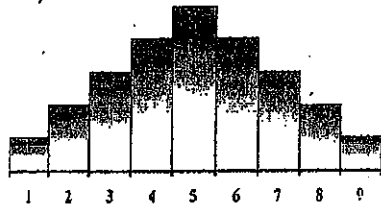
**Less Spread**



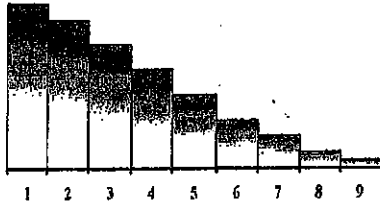
**More Spread**

**SHAPE**

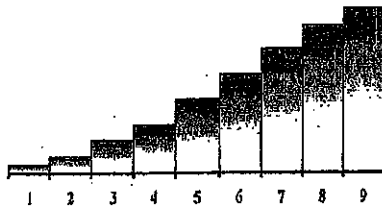
The shape of a distribution is described by symmetry, number of peaks; direction of skew, or uniformity



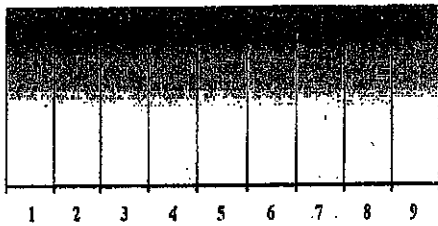
**Symmetric, Unimodal, Bell-shaped**



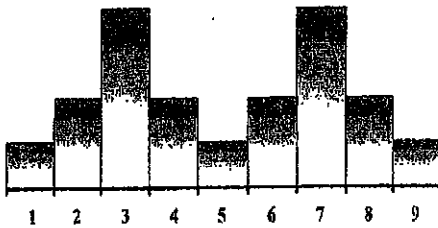
**Skewed Right**



**Skewed Left**



**Uniform**



**Symmetric, Bimodal**



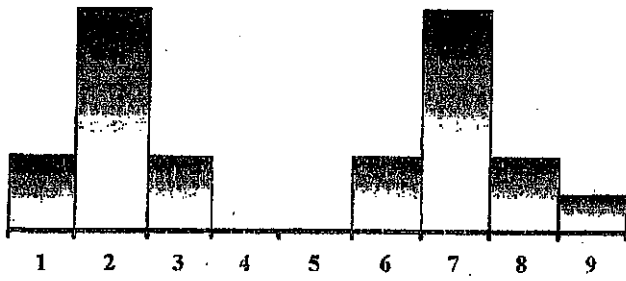
**Non-Symmetric, bimodal**

(14)

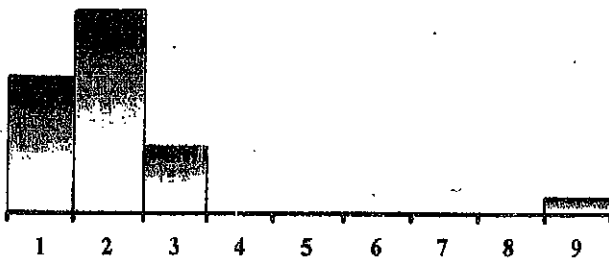
---

### UNUSUAL FEATURES

Sometimes, statisticians refer to unusual features in a set of data. The two most common unusual features are gaps and outliers.



Gap



Outlier

## 8.3 After School Activity

### A Develop Understanding Task



©2012 <http://flickr/y/ghu07>

#### Part I

Rashid is in charge of determining the upcoming after school activity. To determine the type of activity, Rashid asked several students whether they prefer to have a dance or play a game of soccer. As Rashid collected preferences, he organized the data in the following two-way frequency table:

	Girls	Boys	Total
Soccer	14	40	54
Dance	46	6	52
Total	60	46	106

Rashid is feeling unsure of the activity he should choose based on the data he has collected and is asking for help. To better understand how the data is displayed, it is useful to know that the outer numbers, located in the margins of the table, represent the total frequency for each row or column of corresponding values and are called *marginal frequencies*. Values that are part of the 'inner' body of the table are created by the intersection of information from the column and the row and they are called the *joint frequencies*. Using the data in the table, construct a viable argument and explain to Rashid which after school event he should choose.



Part II: Two way frequency tables allow us to organize categorical data in order to draw conclusions. For each set of data below, create a frequency table. When each frequency table is complete, write three sentences about observations of the data, including any trends or associations in the data.

**Data set 1:** There are 45 total students who like to read books. Of those students, 12 of them like non-fiction and the rest like fiction. Four girls like non-fiction. Twenty boys like fiction.

	Fiction	Nonfiction	Total
Boys			
Girls			
Total			

Observation 1:

Observation 2:

Observation 3:

**Data set 2:** 35 seventh graders and 41 eighth graders completed a survey about the amount of time they spend on homework each night. 50 students said they spent more than an hour. 12 eighth graders said they spend less than an hour each night.

			Total
More than one hour			
Less than one hour			
Total			

Observation 1:

Observation 2:

Observation 3:

Name: \_\_\_\_\_ Date: \_\_\_\_\_

**Task – Public Opinions**

**MCC9-12.S.ID.5** Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data.

A public opinion survey explored the relationship between age and support for increasing the minimum wage. The results are found in the following two-way frequency table.

	For	Against	No Opinion	TOTAL
Ages 21-40	25	20	5	50
Ages 41-60	30	30	15	75
Over 60	50	20	5	75
TOTAL	105	70	25	200

**Frequency Count**

1. In the 41 to 60 age group, what percentage supports increasing the minimum wage? Explain how you arrived at your percentage. What type of probability is this? Joint, marginal, or conditional?
2. Out of the people that have no opinion, what percentage is over 60 years old?
3. What are the marginal frequencies?
4. What are the joint frequencies?
5. Why are joint and marginal frequencies important when describing trends or associations in data? Do you see any significant trends when looking at the frequencies?

Name:

## Modeling Data 8.3

Set

Topic: Two-way frequency tables.

Here is the data from Mr. Austin's ten-point quiz. Students needed to score a 6 or better to pass the quiz.

1 <sup>st</sup> Period Math	2 <sup>nd</sup> Period Math	3 <sup>rd</sup> Period Math
6, 4, 3, 7, 5,	3, 3, 8, 6, 6,	9, 8, 10, 5, 9,
9, 5, 4, 6, 6,	9, 5, 8, 5, 3,	7, 8, 9, 8, 3,
8, 5, 7, 3, 6,	5, 5, 7, 5, 7	8, 10, 8, 7, 5
2, 8, 7, 10, 9		

3. Make a two-way frequency table showing how many students passed the quiz and how many failed in each class.

	1 <sup>st</sup> Period	2 <sup>nd</sup> Period	3 <sup>rd</sup> Period	Total
Passed				
Failed				
Total				

4. Use a colored pencil to lightly shade the cells containing the *joint frequency* numbers in the table. The un-shaded numbers are the *marginal frequencies*. (Use these terms to answer the following questions.)

5. If Mr. Austin wanted to see how many students in all 3 classes combined passed the quiz, where would he look?

6. If Mr. Austin wanted to write a ratio of the number of passing students compared to the number of failing students for each class, where would he find the numbers he would need to do this?

7. Make a two-way frequency table that gives the *relative frequencies* of the quiz scores for each class.

	1 <sup>st</sup> Period	2 <sup>nd</sup> Period	3 <sup>rd</sup> Period	Total
Passed				
Failed				
Total				

© 2012 Mathematics Vision Project| MVP



In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.

Name:

## Modeling Data | 8.3

Go

8. Sophie surveyed all of the 6<sup>th</sup> grade students at Reagan Elementary School to find out which TV Network was their favorite. She thought that it would be important to know whether the respondent was a boy or a girl so she recorded her information this way.

Animal Planet	Cartoon Network	Disney	Nickelodeon
GGBBBB BGBBBG GGBB BBBB	BBBBBB BBGGBBG BGBGGGG	GGGGGBBBBB GBGBGG BBGGGG GGBBGGGG	BBBGGGGGGG GGGGGG GGGGGGGGGGGGGG BGGGGGGG

Sophie planned to use her data to answer the following questions:

- I. Are there more girls or boys in the 6<sup>th</sup> grade?
- II. Which network was the boys' favorite?
- III. Was there a network that was favored by more than 50% of one gender?

But when she looked at her chart, she realized that the data wasn't telling her what she wanted to know. Her teacher suggested that her data would be easier to analyze if she could organize it into a two-way frequency chart. Help Sophie out by putting the frequencies into the correct cells.

<i>Favorite TV Networks</i>	<i>Girls</i>	<i>Boys</i>	<i>Totals</i>
Animal Planet			
Cartoon Network			
Disney			
Nickelodeon			
<i>Totals</i>			

Now that Sophie has her data organized, use the two-way frequency chart to answer her 3 questions.

- a. Are there more girls or boys in the 6<sup>th</sup> grade?
- b. Which network was the boys' favorite?
- c. Was there a network that was favored by more than 50% of one gender?

Need Help? Check out these related links:

<http://stattrek.com/statistics/two-way-table.aspx>

© 2012 Mathematics Vision Project | MVP

In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.



## 8.4 Relative Frequency

### *A Solidify Understanding Task*

Rachel is thinking about the data she and her mom collected for the average number of texts a person sends each day and started thinking that perhaps a two-way table of the data they collected would help convince her mom that she does not send an excessive amount of texts for a teenager. The table separates each data point by age (teenager and adult) and by the average number of texts sent (more than 100 per day or less than 100 per day).

	Average is more than 100 texts sent per day	Average is less than 100 texts sent per day	Total
Teenager	20	4	24
Adult	2	22	24
Total	22	26	48

Write two observation statements of this two way table.

To further provide evidence, Rachel decided to do some research. She found that only 43% of people with phones send over 100 texts per day. She was disappointed that the data did not support her case and confused because it did not seem to match what she found in her survey. What questions do these statistic raise for you? What data should Rachel look for to support her case?

After looking more closely at the data, Rachel found other percentages within the same data that seemed more accurate with the data she collected from her teenage friends. How might Rachel use the data in the two way table to find percentages that would be useful for her case?

Part II: Once Rachel realized there are a lot of ways to look at a set of data in a two way table, she was self-motivated to learn about *relative frequency tables* and conditional frequencies. When the data is written as a percent, this is called a *relative frequency table*. In this situation, the 'inner' values represent a percent and are called *conditional frequencies*. The conditional values in a *relative frequency table* can be calculated as percentages of one of the following:

- the whole table (relative frequency of table)
- the rows (relative frequency of rows)
- the columns (relative frequency of column)



Name:

## Modeling Data 8.4

Set

Topic: Relative Frequency Tables

For each two-way table below, create the indicated relative frequency table and also provide two observations with regard to the data.

7. This table represents survey results from a sample of students regarding mode of transportation to and from school.

	Walk	Bike	Car Pool	Bus	Total
Boys	37	47	27	122	233
Girls	38	22	53	79	192
Total	75	69	80	201	425

Create the *relative frequency of row table*. Then provide two observation statements.

	Walk	Bike	Car Pool	Bus	Total
Boys					
Girls					
Total	100%	100%	100%	100%	100%

8. The two-way table contains survey data regarding family size and pet ownership.

	No Pets	Own one Pet	More than one pet	Total
Families of 4 or less	35	52	85	172
Families of 5 or more	15	18	10	43
Total	50	70	95	215

Create the *relative frequency of column table*. Then provide two observation statements.

	No Pets	Own one Pet	More than one pet	Total
Families of 4 or less				100%
Families of 5 or more				100%
Total				100%

© 2012 Mathematics Vision Project | MVP

In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.

Name:

## Modeling Data 8.4

9. The two-way table below contains survey data about boys and girls shoes.

	Athletic shoes	Boots	Dress Shoe	Total
Girls	21	35	60	116
Boys	50	16	10	76
Total	71	51	70	192

Create the *relative frequency of whole table*. Then provide two observation statements.

	Athletic shoes	Boots	Dress Shoe	Total
Girls				
Boys				
Total				100%

Go

Topic: One variable statistical measures and comparisons

**For each set of data determine the mean, median, mode and range. Then create either a box-and-whisker plot or a histogram.**

10. 23, 24, 25, 20, 25, 29, 24, 25, 30

11. 20, 24, 10, 35, 25, 29, 24, 25, 33

12. How do the data sets in problems 10 and 11 compare to one another?

13. 2, 3, 4, 5, 3, 4, 7, 4, 4

14. 1, 1, 3, 5, 5, 10, 5, 1, 14

15. How do the data sets in problems 13 and 14 compare to one another?

© 2012 Mathematics Vision Project | MVP

In partnership with the Utah State Office of Education

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.