Table of Contents

Chapter 7: Probability and Statistics (3 weeks)

UTAH CORE Standards
Probability and Statistics:

Use random sampling to draw inferences about a population.

1. Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.**7.SP.1**

2. Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions. *For example, estimate the mean word length in a book by randomly sampling words from the book; predict the winner of a school election based on randomly sampled survey data. Gauge how far off the estimate or prediction might be*. **7.SP.2**

Draw informal comparative inferences about two populations.

3. Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability. *For example, the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable*. **7.SP.3**

4. Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. *For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book*. **7.SP.4**

Investigate chance processes and develop, use, and evaluate probability models.

5. Understand that the probability of a chance event is a number between 0 and 1 that expresses the likelihood of the event occurring. Larger numbers indicate greater likelihood. A probability near 0 indicates an unlikely event, a probability around 1/2 indicates an event that is neither unlikely nor likely, and a probability near 1 indicates a likely event. **7.SP.5**

6. Approximate the probability of a chance event by collecting data on the chance process that produces it and observing its long-run relative frequency, and predict the approximate relative frequency given the probability. *For example, when rolling a number cube 600 times, predict that a 3 or 6 would be rolled roughly 200 times, but probably not exactly 200 times*. **7.SP.6**

7. Develop a probability model and use it to find probabilities of events. Compare probabilities from a model to observed frequencies; if the agreement is not good, explain possible sources of the discrepancy. **7.SP.7**
   a. Develop a uniform probability model by assigning equal probability to all outcomes, and use the model to determine probabilities of events. *For example, if a student is selected at random from a class, find the probability that Jane will be selected and the probability that a girl will be selected*. **7.SP.7a**
   b. Develop a probability model (which may not be uniform) by observing frequencies in data generated from a chance process. *For example, find the approximate probability that a spinning penny will land heads up or that a tossed paper cup will land open-end down. Do the outcomes for the spinning penny appear to be equally likely based on the observed frequencies?* 7.**SP.7b**

8. Find probabilities of compound events using organized lists, tables, tree diagrams, and simulation. **7.SP.8**
   a. Understand that, just as with simple events, the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs. **7.SP.8a**

b. Represent sample spaces for compound events using methods such as organized lists, tables and tree diagrams. For an event described in everyday language (e.g., "rolling double sixes"), identify the outcomes in the sample space which compose the event. **7.SP.8b**

c. Design and use a simulation to generate frequencies for compound events. *For example, use random digits as a simulation tool to approximate the answer to the question: If 40% of donors have type A blood, what is the probability that it will take at least 4 donors to find one with type A blood?* **7.SP.8c**

Chapter 7 Summary:

Throughout this chapter students engage in a variety of activities: gathering data, creating plots, and making comparisons between data sets. Activities are designed to help students move from experiences to general speculations about probability and number.

Section 1 begins with an exploration of basic probability and notation, using objects such as number cube (dice) and cards. Students will develop modeling strategies to make sense of different contexts and then move to generalizations. In order to perform the necessary probability calculations, students work with fraction and decimal equivalents. These exercises should strengthen students' abilities with rational number operations. Some probabilities aren't known, but can be estimated by repeating a trial many times, thus estimating the probability from a large number of trials. This is known as the Law of Large Numbers, and will be explored by tossing a Hershey's Kiss many times and calculating the proportion of times the Kiss lands on its base.

Section 2 investigates the basics of gathering samples randomly in order to learn about characteristics of populations, in other words, the basics of inferential statistics. Typically, population values are not knowable because most populations are too large or difficult to measure. "Inferential statistics" means that samples from the population are collected, and then analyzed in order to make judgments about the population. The key to obtaining samples that represent the population is to select samples randomly. Students will gather samples from real and pretend populations, plot the data, perform calculations on the sample results, and then use the information from the samples to make decisions about characteristics of the population.

Section 3 uses inferential statistics to compare two or more populations. In this section, students use data from existing samples and also gather their own data. They compare plots from the different populations, and then make comparisons of center and spread of the populations, through both calculations and visual comparisons.

Terms and phrases used in this chapter are informally explained below.

VOCABULARY:

random sample – a set of data that is chosen in such a way that each member of the population has an equal probability of being selected

population – the set of possibilities for which data can be selected

independent events – events that are not affected by each other

compound events – an event made up of two or more independent events

expected value – the average value of repeated observations in a replicated experiment

frequency – the number of times that a particular value occurs in an observation

probability – the chance or likelihood that an event will occur, expressed from a scale from 0 (impossible) to 1 (certain)

relative frequency – the ratio of the frequency of an event in an experiment to the total frequency

Law of Large Numbers – the long run relative frequency of an experiment, based on a large number of trials

sample – a subset of a population collected by a defined procedure for the purpose of making inferences from the sample to the population

simulation – an experiment that models a real-life situation

probability model – a mathematical representation of a random phenomenon that includes listing the sample space and the probability of each element in the sample space

uniform probability model – when all of the outcomes of a probability model are equally likely

CONNECTIONS TO CONTENT:

Prior Knowledge

Students should be familiar with the following content from 6th grade:

- Understands that a set of data has a distribution that can be described by its center, spread, and overall shape. **6.SP.2**
- Displays numerical data in plots on a number line, dot plots, histograms, and box plots. **6.SP.4**
- Gives quantitative measures of center (median and/or mean) and variability (IQR and/or mean absolute deviation) **6.SP.5c**
- Describes any overall patterns of data and any striking deviations from the overall pattern. **6.SP.5c**
- Relates the choice of measure of center and variability to the shape of the data distribution and context. **6.SP.5d**

Chapter 7 begins by reviewing standard 7.SP.5, basic probability content that was covered in Chapter 1.

Future Knowledge

This unit introduces the importance of fairness in random sampling, and of using samples to draw inferences about populations. Some of the statistical tools used in 6th grade will be practiced and expanded upon as students continue to work with measures of center and spread to make comparisons between populations. Students will investigate chance processes as they develop, use, and evaluate probability models. Compound events will be explored through simulation, and by multiple representations such as tables, lists, and tree diagrams.

The eighth grade statistical curriculum will focus on scatter plots and bivariate measurement data. Bivariate data is also explored in Secondary Math I, however, Secondary Math I, II, & III statistics standards return to exploration of center and spread, random probability calculations, sampling and inference.

MATHEMATICAL PRACTICE STANDARDS (emphasized):

	Make sense of problems and persevere in solving them.	Students will make sense of probability calculations by connecting rational numbers to probabilities, and creating models to support calculations. Additionally, students will use sense-making skills to compare data sets using measures of center and spread.
	Reason abstractly and quantitatively	Students are able to utilize the mathematics necessary to solve simple probability problems using both ratios and percents, and interpret data using appropriate measures of center and spread.
	Construct viable arguments and critique the reasoning of others	Students are able to assess the reasonableness of their answers and will solve problems in a variety of ways, where they will be able to discuss and validate their own approaches and solutions.
	Model with mathematics	Students will use multiple representations to model probability problems and create appropriate graphical representations for data.
	Attend to precision	Students will identify whether or not their answer makes sense (e.g. probability values less than 0 or greater than 1 are not valid, measures of center and spread should be reasonable for the data).
	Look for and make use of structure	Students are able to recognize the key phrases of compound probability models and use of diagrams or tables to assist with calculations and data analysis.
	Use appropriate tools strategically	Students demonstrate their ability to select and use the most appropriate tool(s), such as diagrams, tables, lists, box plots, dot plots, etc., while solving real-life word problems.
	Look for and express regularity in repeated reasoning	Students look for structure and patterns in real-life word problems, which will help them identify a solution strategy.

7.0 Anchor Problem: The Teacher Always Wins

The Teacher Always Wins Game (to be introduced by the teacher)
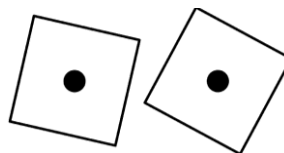
Why does the teacher always seem to win? Is it certain that the teacher will win? After the introduction of this activity, your job is to determine the answers to these questions, and see if you can discover the secret to the game.

Section 7.1: Probability Models to Analyze Real Data, Make Predictions

Section Overview:

This section starts with a review of concepts from Chapter 1 section 1 and then extends to a more thorough look at probability models. A complete probability model includes a sample space that lists all possible outcomes, including the probability of each outcome. The sum of the probabilities from the model is always 1. A uniform probability model will have relative frequency probabilities that are equivalent. A probability model of a chance event (which may or may not be uniform) can be approximated through the collection of data and observing the long-run relative frequencies to approximate the theoretical probabilities. Probability models can be used for predictions and determining likely or unlikely events.

There are multiple representations of how probability models can be displayed. These include, but are not limited to: organized lists (including a list that uses set notation), tables, and tree diagrams.

Students will also consider the ramifications of rounding, what it means to have "independent events," how to create a simulation, and further explore the difference between theoretical probability and real life situations. There will be several exploration activities in the section giving students ample opportunity to discuss ideas.

Concepts and Skills to be Mastered
1. Develop a probability model and use it to find probabilities of events.
2. Compare probabilities from a model to observed frequencies; if the agreement is not good, explain possible sources of the discrepancy.
3. Understand that, just as with simple events, the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs.
4. Represent sample spaces for compound events using methods such as organized lists, tables, and tree diagrams.

7.1a Class Activity: The Horse Race Game Revisited—Probability Basics

Note: You may wish to use ratio, decimal or percents to discuss probability and compute predictions. If you did not do 1.1.b, you should do it BEFORE this activity. Review with students the difference between saying "there is a 1 in 3 chance of drawing a red marble from this bag of red and blue marbles" and "there is 1 red marble for every 3 blue marbles in the bag."

1. Review from Chapter 1 (1.1.b): in the Horse Race Game you predicted which horse (#2-12) would win the race. The winning horse was determined by tossing two dice and observing the sum of the die.   Fill in the table below to find the *possible* sums of two dice.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

In general, the "#7 horse" wins. There is a 1/6 probability of rolling a sum of 7 while all other rolls (sums) have probabilities less than 1/6. See 1.1b for more information.

Use the table to answer the following questions. Some of the questions may review Chapter 1 content.

2. What is the number of the horse that is most likely to win? Explain how you know.
   Horse #7 Seven is the sum that occurs most often.

3. How many times did that horse's number occur in the table?   6 times

4. What is the number of the horse (or horses) that is/are the least likely to win? Both horses #2 and #12
   Remind students why there is no "1 horse."

5. How many times did that horse(s) number occur in the table? _____1 each

6. How many total outcomes are there altogether on the table? _____ 36 outcomes

RECALL: Probability is written as a part-to-whole ratio of possible outcomes to the number of total outcomes. For example, the chances that horse #3 will win is: two *possible* ways to win out of thirty-six *total* possibilities (or 2/36.)

7. What is the probability that horse #8 will win? _____ 5/36

Note that the word "probability" is used here. Talk about other words such as "chance" that are sometimes used instead. The word "odds" is NOT correct. Odds and probability are related but distinct numeric representations of situations.

The table shows the probabilities for each horse winning. Recall, this is called the **theoretical probability** of winning. Remind students that we didn't actually race the horses, but we can estimate the probabilities that each one might win using the theoretical probability.
Mathematical Notation: the mathematical shorthand way of writing a probability looks like this:
P(horse #4 wins) = 3/36 OR P(4) = 3/36 (we can also write this as a reduced fraction, decimal or percent.)

8. Fill in the table below with the theoretical probability for each horse to win. Write the values both as a fraction and as a percent. Round the percents to the nearest whole number.

Number	2	3	4	5	6	7	8	9	10	11	12
Probability as a fraction	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
Probability as a percent	3%	6%	8%	11%	14%	17%	14%	11%	8%	6%	3%

This is a good opportunity to review fractions. Note that it is perfectly acceptable in probability calculations to leave the fraction in the original form because it provides information about the size of the original sample and the number of outcomes of each type.

9. Add up all the fractions. What is the total? The total is 1.
Vocabulary: explicitly discuss that the table showing the outcomes (horses #2-12) along with probabilities that add to 1 represent a **Probability Model**. The table shown above is one representation of a probability model. Other representations include ordered lists and tree diagrams. The sum of all probabilities in a probability model will be 1, or if in percents, 100%. If the values don't add to 1 (or 100%), then you're missing something and/or there is some problem with the way you're calculating.

10. Add up all of the percents in the table. What is the total? 101%. Let the students think about this. Discuss why the sum is not 100%. Discuss how *rounding error* occurs.

11. If there are 200 races (200 rolls of the dice), how often would you predict horse #7 would win? Show all your work and explain your reasoning.
(200)(6/36) = 1200/36 or 100/3 or $33.\overline{3}$ races (not a whole number), or using rounded percents: (200)(0.17) = 34 races (a whole number).

12. Suppose the horses race…
   a. …500 times, what is your prediction for how many times horse #7 will win? Show your calculations.
   (500)(6/36) = 83.33, or using the less accurate rounded percents: (500)(0.17) = 85.

   b. …1000 times, what is your prediction for how many times horse #2 will win? …horse #12 will win?
   Each will equal (1000)(1/36) = $27.\overline{7}$, or about 28 times.

   c. Suppose we watched the horses race 500 times. Which of the following values would be the most likely result for horse #5?   11 wins   50 wins*   100 wins   250 wins

   Explain the reasoning for your choice.   50 is the closest to the value you would expect. Students should attend to the fact that winning/losing do not have equal probability. It may be useful to ask students: Why isn't the answer 250? Horse #5 can only either win or lose, so isn't that a 50% chance of winning? Pin down that winning and losing are not equally likely outcomes. In other words, even though there are only two outcomes possible, winning has a 4/36 chance while losing has a 32/36 chance.

7.1a Homework: Probability Problem Solving

M&M Probability (refer to the table below for amounts of colors)

1. The color mix in a large bag of M&Ms is shown in the table below. What is the total number of M&M's in the bag?   The total number of M&Ms = 220

2. Calculate the probability of drawing each of the colors.   Finish the probability model by recording the experimental probability of drawing each color.   Show the probabilities as both a fraction and as a percent.

Color and number	RED 60	GREEN 40	BROWN 45	YELLOW 25	ORANGE 20	BLUE 30
Fractions	60/220 = 3/11					
Percents	27%					

3. If you drew 50 M&M's, one at a time (returning the M&M to the bag each time), how many of each color would you expect, based on the probabilities in the table above?   Put your answers in the table.

Predicted Sampling Estimate for 50 draws					
RED	GREEN	BROWN	YELLOW	ORANGE	BLUE
13.64 ≈ 14					

Remind students that the predicted numbers should add up to 50 and why this may not be the case when rounding.

4. Suppose you went to the store and bought a large bag of M&Ms. From that bag you took a sample of exactly 50 M&Ms and calculated the percent of each color in your sample. Do you think the percents would be the same as in the first table? Why or why not?

> **Probability model** – a mathematical representation of a random phenomenon that includes listing the sample space and the probability of each element in the sample space.

5. The Bag Game

There are three bags of chips: one with 25 red and 5 blue, another with 20 red and 10 blue, and the last with 10 red and 20 blue. You're randomly given one bag. To win the game, you must guess correctly which bag you've been given but you cannot see its contents.

To make your guess you are given three options:

a. Draw 5 chips and guess correctly, win $100.   d. Draw 20 chips and guess correctly, win $25.
b. Draw 10 chips and guess correctly, win $75.   e. Draw 25 chips and guess correctly, win $10.
c. Draw 15 chips and guess correctly, win $50.

Note: to plat the bag game, you draw one chip at a time, record the color, replace the chip, and then repeat.

You want to win as much money as possible. Which option do you choose for guessing which bag you've been given? Why? Be certain to explain all of the probabilities. Discuss with students that there is not a "right" answer. What matters is soundness of the argument. You might want to talk about insurance rates, sports, or other areas where probabilities are considered to make decisions.

6. Rolling Doubles

If TWO dice are rolled 36 times, how many doubles would you expect to see? What is the probability of rolling doubles with two fair die? Students may construct a table similar to the one at the beginning of the activity. Instead of listing each sum, list the dice combinations like (1,6), (2,6) etc.

Spiral Review

1. Write 0.612 as a percent and fraction.   61.2 % 612/1000 or 153/250

2. If 4 gallons of gas cost $14.60, how much does 10 gallons of gas cost? $36.50

3. If you spin the following spinner once, what is the theoretical probability of spinning an L?
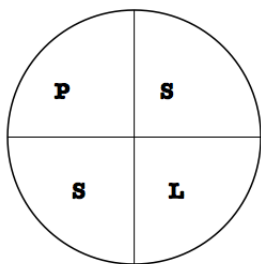
4. A mouse can travel 1.5 miles in ¾ of an hour. At that pace,
    a) how far can it travel in 1 hour? 2 miles
    b) how long does it take it to travel one mile? ½ an hour

Madison is riding her horse around the outside of a circular arena. She knows that 14 laps is ½ mile. What is the diameter of the arena? (Hint: 1 mi = 5280 ft)

7.1b Class Activity: Probability Models

Probability models (like "tree" models) show the outcome of random processes. A probability model includes the following:

- A listing of the sample space (all the possible outcomes.) For example, you might use set notation S = { _ , _ , _ , …}, a tree diagram, a table, etc.
- Probability for each possible event in the sample space. Remember, probabilities always add up to 1.

Talk again to students about notation: set notation for a sample space appears as: $S = \{a, b, c,…\}$. The "S" stands for "sample space". The curly brackets enclose the possible outcomes. Each possible outcome is usually only listed once, even if it occurs more than once.

1. Suppose you are going to toss a coin and see how it lands.
    a. List the sample space using set notation. $S = \{H, T\}$

    b. What is the probability for tossing a head? P(head) =   0.5 or 1/2

    c. What is the probability for tossing a tail? P(tail) = 0.5 or 1/2

    Ask students: what if you were to toss a thumb tack, what are the possible outcomes? You might also (or instead) ask: what if you were to toss a Kleenex box, what are the possible outcomes for how the box will land?
    Follow up either situation by asking if all outcomes are equally likely. The answer is NO. For the thumb tack, it might land up (on the back of the tack), sideways, or on the tip of the tack, but these are not equally likely. Likewise with the Kleenex box. Though there are 6 sides of the box, students will note that it is more likely that the box will fall on to one of the sides with the greatest surface area.

2. Consider the theoretical outcomes for tossing a fair coin 3 times.
    a. What is the sample space? Use set notation. (Hint: there should be 8 outcomes in the sample space.)
       {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

    b. What is the probability of each of the 8 outcomes? 1/8. Discuss: Since heads and tails are equally likely, each of the eight outcomes in the sample space are equally likely.

A probability model for which all outcomes are equally likely (have the same probability) is called a **uniform probability model**.

3. Create a tree diagram to display the sample space for tossing a coin 3 times. **The first branch should have two forks, one for H and one for T. Each of those has two forks (now 4 outcomes), and each of those have two forks (now 8 outcomes).**

			Outcomes	**Probability**

1st Toss	2nd Toss	3rd Toss	Outcomes	Probability
		H	H H H	(1/2)(1/2)(1/2)=1/8
		T	H H T	(1/2)(1/2)(1/2)=1/8
		H	H T H	(1/2)(1/2)(1/2)=1/8
		T	H T T	(1/2)(1/2)(1/2)=1/8
		H	T H H	(1/2)(1/2)(1/2)=1/8
		T	T H T	(1/2)(1/2)(1/2)=1/8
		H	T T H	(1/2)(1/2)(1/2)=1/8
		T	T T T	(1/2)(1/2)(1/2)=1/8
				Total: 1

A tree diagram also helps organize information so that students can list all of the outcomes without missing any of them. Notice that a tree diagram is also handy for calculating probabilities. Since this is a uniform distribution, the chance of being on either fork is equally likely at 1/2. Therefore, the probability of HHH, or P(HHH), is (1/2) · (1/2) · (1/2) = 1/8.

4. Use the list or the tree diagram for 3 coin tosses to fill in the theoretical probability of the following events:

3 heads	3 tails	2 heads and 1 tail (in any order)	2 tails and 1 head (in any order)
1/8 or 0.125	1/8 or 0.125	3/8 or 0.375	3/8 or 0.375

Ask the students to search for "2 heads and 1 tail (in any order)" on the tree diagram.

Ask students if they think the table shown above represent a complete probability model? How do they know? All possible outcomes are shown along with the probabilities, and the probabilities add up to 1.

5. Use all or part of the tree diagram in #4 to calculate the following probabilities:
   Notation: if you see P(TTH), that is the same as writing "the probability of a tail, and a tail, and a head"

$P(H) = 0.5$ or 1/2 $P(T) = 0.5$ or 1/2

$P(HT) = (0.5)(0.5) = 0.25$ or 1/4 $P(TH) = (0.5)(0.5) = 0.25$ or 1/4

$P(HTH) = (0.5)(0.5)(0.5) = 0.125$ or 1/8 $P(TTT) = (0.5)(0.5)(0.5) = 0.125$ or 1/8

Compound Event: an event made up of two or more independent events.

---

6. Fill in the blanks for calculating probabilities for **compound events,** in other words, for two or more events occurring together.
   Suppose we call the first event A, the second event B, the third event C, etc.

   If P(A) = 0.5 and P(B) = 0.5
   then the compound probability P(A and B) = _____ (0.5)(0.5) = 0.25

   If P(A) = 0.5 and P(B) = 0.5 and P(C) = 0.5
   then the compound probability P(A and B and C) = _____ (0.5)(0.5)(0.5) = 0.125

   Using words and symbols, state your conjecture for the general rule for calculating probabilities for independent compound events. In our activity above we wrote P(HTH) as shorthand for P(head the tail the head). To make a general rule, we will use A, B, and C. Thus, a general rule might look like: P(A and B and C) = P(A)·P(B)·P(C) Note again, that this rule is only true for independent events. It is another lesson entirely to consider compound events that are not independent. This topic is addressed in Secondary Math Vocabulary: Independent – the outcome of one event has no effect on the next event. Ask students to describe how the rule for independent compound probabilities relates to the tree diagram. One possible answer: you can find the probability of the end result (such as following the branches leading to HTH) by multiplying the probabilities along each branch. Also discuss explicitly that P(HTH) is different than P(THH) or P(HHT), but that each has the same value. If we want to find the probability of flipping two heads and one tail in three flips, any order, then we find all the ways this is possible (THH, HTH or HHT) and sum each of the three ways it can be achieved: 0.125 + 0.125 + 0.125 or 0.375.

7. Use the rule you found in the prior question to calculate P(HTHHH). (0.5)(0.5)(0.5)(0.5)(0.5) = 0.03125 = 1/32. Again emphasize that you are asking for heads and tails in a specific order in five flips.

8. Suppose that you have an *unfair* coin where the P(H) = 0.8 and the P(T) = 0.2. Compute the following probabilities:
   P(H) = 0.8   P(T) = 0.2

   P(HT) = (0.8)(0.2) = 0.16   P(TH) = (0.2)(0.8) = 0.16

   P(HTH) = (0.8)(0.2)(0.8) = 0.128   P(TTT) = (0.2)(0.2)(0.2) = 0.008

Note: This is not a uniform distribution because the outcomes are not equally likely.

9. Compare the calculations that you used for the fair coin and the unfair coin. How are the calculations similar? How are they different? You use the same method of multiplying the probabilities together, but you change the probabilities to P(H) = 0.8 and P(T) = 0.2

Talk with students about how they might continue to use the tree diagram to calculate compound probabilities for an unfair coin. Help them understand that the tree diagram is still useful, you just have to change the probabilities. The outcomes are no longer equally likely.

7.1b Homework: Probability Models

Suppose you rolled a dice and tossed a coin at the same time.
1. Create a probability model, BOTH a tree model and table, for rolling a die once then tossing a coin once.

2. How many total outcomes are represented by either the tree or table model ? $6 \cdot 2 = 12$

3. What is the sample space for the possible outcomes? List the sample space using set notation.
   $S = \{$ ___ , ___ , ____ ...$\}$ $S = \{$1H, 1T, 2H, 2T, 3H, 3T, 4H, 4T, 5H, 5T, 6H, 6T$\}$. Can be in any order, and can be described using words rather than numerals and letters.

4. What is the probability for each outcome in the sample space? Write the probabilities both as a ratio and as a percent.

5. If you collected experimental data from rolling a die and then tossing a coin, would the calculated probabilities from the experiment match the theoretical probabilities? Why or why not?

Spiral Review

1. Rewrite the following part:part ratio as part:whole ratio.
   a. The ratio of boys to girls in Gabrielle's family is 3:8.   3:11

2. Solve the following proportion equation: $\frac{x}{5} = \frac{4}{10}$ $x = 2$

3. Simplify each.
   a) $-6(-5)$ 30   b) $-10 \cdot 31$ -310

4. Kim had a bag with red, green, purple, yellow and orange marbles. The following table shows what color she drew each time.

Draw	1	2	3	4	5	6	7	8	9	10
Result	red	orange	purple	orange	orange	purple	yellow	green	red	green

   a) Find the experimental probability of drawing a red marble. $\frac{1}{5}$

   b) If there are 100 marbles in the bag, how many of them do you think are red? Justify your answer.
   20

5. If ∠N is vertical to ∠M, and $m\angle M = 98°$ and $m\angle N = (6x + 2)°$, the x must be _____.

7.1c Class Activity: Rolling Along

Can you roll your tongue? Some people can roll their tongue, some cannot.
Approximately 1 out of every 3 people *cannot* roll their tongue.
Depending on the source, the estimated proportion of people in the population who cannot roll their tongues is between 19-35%.
Consider: Ria is doing a survey on the number of people with different genetic traits. She asks people, one at a time, if they can roll their tongue. Ria was surprised that she asked 5 people before she found someone who wasn't able to roll their tongue. **Does this mean the statement "approximately 1 out of 3 people cannot roll their tongue" must be false? Is it unusual that after surveying 5 people she did not find anyone who could not roll their tongue?**

To answer this question we can do a simulation of Ria's experiment.

Simulation – an experiment that models a real-life situation

Select a method to simulate a 1 out of 3 chance (die, slips of paper, software, etc.). Run the simulation until you get the "1 out of 3" chance you're looking for. For example, there is a 1 out of 3 chance of rolling a 1 or 2 with a six sided die. *One simulation is the number of times it takes to roll a 1 or 2 with a die*. Record a tally mark under the number of times it takes to get the 1 or 2 in the table below. Run the simulation 20 times recording your result each time. Once you've done your 20 simulations, compile your results with two other people so that you have 60 total simulations. Record the data in the table**.**

Number of attempts	1	2	3	4	5	6	7	8	9	10
Record a tally	7	4	3	2	1.3	0.9	0.6	0.4	0.3	0.2
Combined Results (60 simulations)	20	13	9	6	4	3	2	1	0.8	0.5
% out of 60 simulations	20/60 33.33%	13/60 22.67%	9/60 15%	6/60 10%	4/60 6.67%	3/60 5%	2/60 3.33%	1/60 1.67%	0.8/60 1.33%	0.5/60 0.83%

1) Based on the combined trials, calculate the probability that it would take 5 or more attempts.
   Theoretically, P(5 or more attempts) = 0.1975. When students combine their results to obtain 60 total trials, they should get a value of around 11 where it took 5 or more attempts to get a success, although student results will vary. As an example, this table shows expected values for up to 10 attempts. Summing the values found in the table for 5 or more attempts: 4 + 3 + 2 + 1 + 0.8 + 0.5 = 11.3. Thus students would expect around 11 of the 60 trials to take five or more attempts.
2) Were Ria's results unusual? Write a paragraph summarizing your conclusion, based on the simulation. No, Ria's results are not terribly unusual. The simulation shows that results like this could happen in slightly 2-3 times out of 20, and about 11 times out of the combined 60 attempts. A discussion of the variability of the students' results is an important concept in statistics, and in simulation.

3) Suppose the ratio of left handed people to right handed people is 1:10. Create a simulation for the number of trials if takes to get a left handed person.

4) Suppose that 60% of students choose chocolate ice cream, 30% choose vanilla ice cream, and 10% choose strawberry ice cream. Create a simulation for the number of trials it takes to get a student to choose chocolate, vanilla and strawberry ice cream.

> **Law of Large Numbers** – the long run relative frequency of an experiment, based on a large number of trials.

7.1c Homework: Finding Probability

The colors of M&Ms in a large bag are distributed according to the probabilities shown in the table:

Color	Brown	Red	Yellow	Green	Orange	Blue
Probability	0.25	0.25	0.20	0.10	0.10	?

1. Finish the table above by finding P(blue).

2. Suppose you draw an M&M out of the bag and record the color. List the sample space using set notation.
$S =$

3. What is required in order to have a complete probability model?
   All of the possible outcomes listed, and probability of each outcome, and the probabilities adding up to 1.

Compound Probabilities of M&M Colors
4. Compute the following theoretical probabilities. Use the probabilities from the M&M table given above.

P(red and yellow) = (0.2)(0.2) = 0.04

P(brown, orange) =

P(3 blues in a row) =

Simulation
5. Your favorite M&M's are red, so you want to create a simulation for modeling the drawing of red M&M's from a bag with the color probabilities as listed above. Describe a simulation. Remember: you are only trying to simulate drawing a red M&M.
   Hint: From the table above, P(red) = 0.25. Theoretically in four tries, one will be red.

1. If you flip a penny three times, what is the probability of getting two tails and one head in any order?

2. The scale factor $\triangle GEL$ to $\triangle HOP$ is $\dfrac{1}{6}$. If $\overline{OP}$ is 30, what is the length of $\overline{EL}$?

3. Daniel got 4 out of every 5 questions correct on a recent multiple choice test. If he got 64 questions correct, how many did he miss?   80 - 64 = 16   Daniel missed 16 questions.

4. Find the sum or difference for each:

    a. $\dfrac{5}{3} - \dfrac{3}{4}$ $\dfrac{11}{12}$

    b. $-\dfrac{2}{3} + \dfrac{1}{4}$ $-\dfrac{5}{12}$

5. Paul left a $25 tip for the waiter at a restaurant. If the tip was 25% of the bill, how much was the bill?
    $100

7.1d Class Activity: More Models and Probability
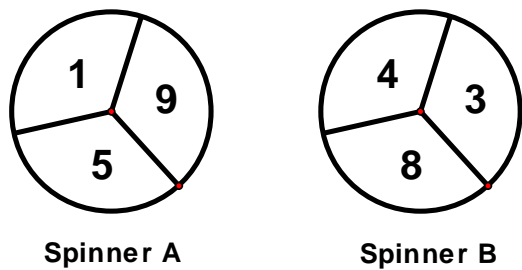
Spinner A   Spinner B

Win the Spin!
Determine the probability that Player 1 wins the spin (highest number wins). Player 1 uses spinner A and Player 2 uses spinner B. Assume that the areas on each spinner are equal in size.

1. Create a probability model for the outcomes of the "Win the Spin" game, using a tree diagram. Students

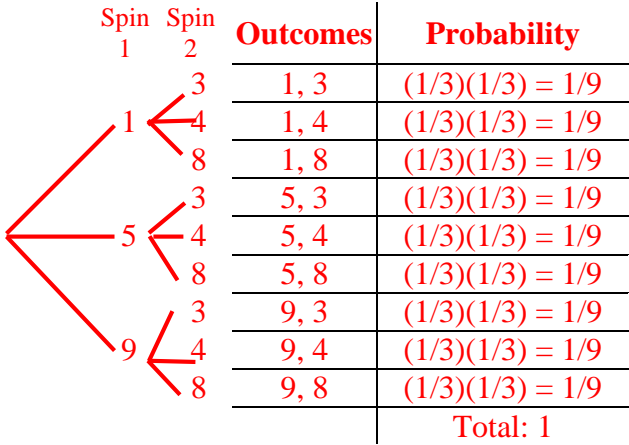|	Spin 1	Spin 2	Outcomes	Probability
	1	3	1, 3	(1/3)(1/3) = 1/9
		4	1, 4	(1/3)(1/3) = 1/9
		8	1, 8	(1/3)(1/3) = 1/9
	5	3	5, 3	(1/3)(1/3) = 1/9
		4	5, 4	(1/3)(1/3) = 1/9
		8	5, 8	(1/3)(1/3) = 1/9
	9	3	9, 3	(1/3)(1/3) = 1/9
		4	9, 4	(1/3)(1/3) = 1/9
		8	9, 8	(1/3)(1/3) = 1/9
				Total: 1

2. How many possible outcomes are there? How do you know? 3 • 3 = 9 outcomes or students could count the outcomes from the tree diagram.

3. What is the probability of each of the outcomes? How do you know? 1/9. A possible answer: the probabilities are equal because the areas of each outcome on the spinner are equal at 1/3. Since the game requires each of the two spinners to be used once, then each of the outcomes from the two spins have a probability of (1/3)(1/3) = 1/9.

4. What is the probability that Player 1 will win? How do you know? Player 1 wins 5 times out of the 9 outcomes, or 5/9.   Refer to the outcomes on the probability model.

5. What is the probability that Player 2 will win? How do you know? Player 2 wins 4/9 times.

6. What is the probability of a draw (tie)? How do you know? P(tie) = 0 There are no ties, because the spinners have different numbers. Make note that the probability of an impossible event is 0. Probabilities have values between 0 and 1.

7. Is this a fair game? Why or why not? No, Player 1 has a higher probability (5/9) of winning.

8. If you were to play the game, your outcome would not necessarily match the probabilities above. Explain why this is true. **The students will probably mention that there is always variability.**

Odd or Even Game

For a different game Player 1 and 2 each spin once (Spinner A, then B) and add the numbers. If the sum is odd, then Player 1 gets a point. If the sum is even, then Player 2 gets a point. Review with students: odd number + odd number = even number; odd number + even number = odd number; even number + even number = even number; and even number + odd number = odd number.

9. Create a probability model for the outcomes of the "Odd or Even" game, using a tree diagram.

10. Use the tree diagram to figure out if the game is fair or not. Explain. **The game is not fair. P(even) = 3/9, so Player 1 only wins 3 times out of 9.**

11. Use the rule for calculating compound probability to calculate the probabilities of the different combinations of the spins for Players 1 and 2. Show all your work.

P(odd, even) = (3/3)(2/3) = 6/9. Verify with the tree diagram. Player 2 will win these.

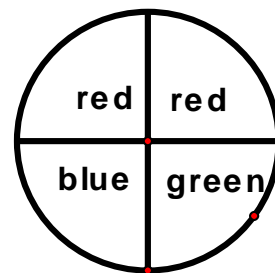P(odd, odd) = (3/3)(1/3) = 3/9. Verify with the tree diagram. Player 1 will win these.

P(even, odd) = 0(1/3) = 0. Verify with the tree diagram.

P(even, even) = 0(2/3) = 0. Verify with the tree diagram.

7.1d Homework: Probability Models and Spinner Games

1. The spinner at right is spun twice. List the sample space for the possible outcomes from two spins. Use set notation. (Hint: there are 9 outcomes)
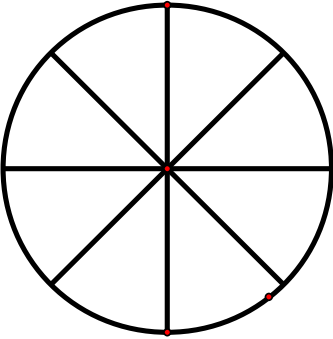
   $S =$
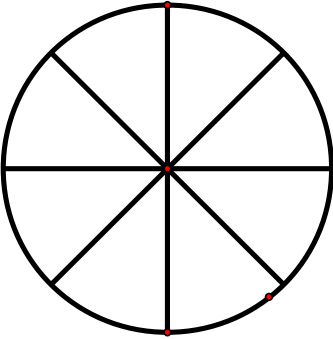
2. Are all the outcomes in the sample space equally likely? Why or why not?

3. How might you figure out the number of outcomes without making a list or diagram? Explain.
   The first spin has 3 different outcomes. Each one of those has 3 different outcomes paired with it for the second spin. So 3 × 3 = 9. This may cause confusion for some students. Discuss how the right red and then blue is the same outcome as the left red and then the blue. i.e. it is P(red, blue). However, it is different than P(blue, red) and that outcome can happen by first spinning a blue and then the red on the right or left. It might help students to think about the entire top half of the spinner as red rather than thinking about it as two different areas.
4. Create a probability model for the outcomes for spinning the spinner twice (organized list, tree diagram, or table) to show all possible outcomes and probabilities from two spins of the spinner.

5. Fill in the spaces below and make a conjecture about a rule for the number of possible outcomes for compound events.

   If there are 3 possible outcomes in the first event, and 2 possible outcomes in the second event, then there will be _____ possible outcomes in the compound event. 2 · 3 = 6

   If there are 5 possible outcomes in the first event, and 3 possible outcomes in the second event, then there will be _____ possible outcomes in the compound event.

   If there are *"a"* possible outcomes in the first event, and *"b"* possible outcomes in the second event, then there will be _____ possible outcomes in the compound event.

7WB 7 - 24

6. Examine the model you created above to determine these probabilities.
   a. P (red, red) = (1/2)(1/2) = 1/4

   b. P (one red and one green, *in any order*)

   c. P (blue, red) Note: blue must come first.

7. Create a probability game for each spinner. Design the spinners to make one game fair and the other unfair. Write the rules to tell how to play each game.

Game 1	Game 2
Game 1 Rules	Game 2 Rules

8. Explain the probability of winning each game and why the game is fair or unfair.

Extra Challenge: Spinners for Math Day

Howard is in charge of the Spinner Game for the Math Fair. There will be about 300 people at the fair and he believes everyone will buy a ticket to play the Spinner Game. The school wants to raise money for some math software. Spinner tickets cost $1. Winners of the spinner game will be given cash prizes. Hal wants to make $100 profit from the game.

Design a plan that should net Hal $100 from the Spinner Game. Be sure to show the spinners you would recommend and the rules you think would work. Explain why your spinners and rules make sense for this context.

Spiral Review

1. Simplify each:

   a. $-6(-5)$ 30

   b. $-16(-3)$ 48

   c. $-10 \cdot 31$ -310

   d. $-89 + (-6)$ -9

2. Kelsey puts each letter of her name on a piece of paper. What is the probability that she will draw a K and an E in any order?

3. Order the following rational numbers from least to greatest. $\frac{12}{3}, -4.5, -\frac{14}{3}, -0.94,$

$$-\frac{14}{3}, -4.5, -0.94, \frac{12}{3}$$

4. Matthew wants a bigger cage for his bearded dragon. He wants to length to be 2 inches less than twice the width. If the perimeter of the cage should be 104 in, what dimensions should the cage be?

5. Use the diagram at the left to find the angle measures.

$2x$-9

x-3
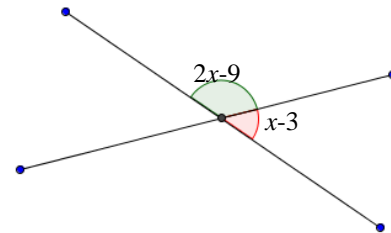
Long run relative frequency: the probability of an outcome obtained after many trials
Variable (the verb, not the noun): not consistent or having a fixed pattern; liable to change
Experimental Probability: the ratio of the number favorable outcomes to the total number of trials, from an actual sequence of experiments
Theoretical Probability: the probability that a certain outcome will occur, as determined through reasoning or calculation

7.1e Class Activity: Probability of a Kiss

When you toss a coin, it will either land heads or tails. That isn't very interesting. But suppose you toss a Hershey's Kiss in the air, and then observe how it lands. That is much more interesting.

The sample space for tossing a Kiss has two possible outcomes: S = {base, side}. "Base" means the Kiss landed on the flat base, and "side" means the Kiss landed on its side.

What is the probability for each outcome? We don't know the answer. First we must do an experiment, and then calculate the experimental probability that the Kiss will land on its base, P(B).

1. Define "long run relative frequency". The probability of an outcome obtained after many trials.

2. Make a guess for the probability that a Kiss tossed in the air will land on its base:
    P(B) = Student answers will vary, most students might think the Kiss will land more frequently on the base, but actually it will land more frequently on the side. There is more surface area on the side than the base. In past trials done by students, the probability of landing on the base was always less than 0.5, sometimes about 0.3

3. Record the results of your experiment in the table after each toss and calculate the experimental probability after each trial. Notice that the probabilities are calculated from:
    (Base running total)/(trial number).

4. Make a plot of your results by plotting the trial number against the experimental probability. Connect the points from trial 1, to trial 2, to trial 3,…and end with trial 30.

Hint: Have students fill in the outcomes column then complete the running total and experimental probability.

Trial Number	Outcome (B or S)	# of times on Base (running total)	Experimental Probability
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			

20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			

Ex 1	B	1	1:1
Ex 2	S	1	1:2
Ex 3	B	2	2:3

5. What is the experimental probability of a Kiss landing on its base after 30 trials for your experiment? P(B) =

.

6. Compare the value for the experimental probability at Trial 2, compared to Trial 20. Which value was closer to the final experimental value you found at Trial 30?

7. Examine the appearance of the plot. Why is the plot so variable at the beginning compared to at the end?

8. Why is it important to perform many trials in an experiment, and not just a few?

Probability of a Kiss: Graph your results below. Draw lines between the points when you are finished.

Experimental Probability

Trial Number
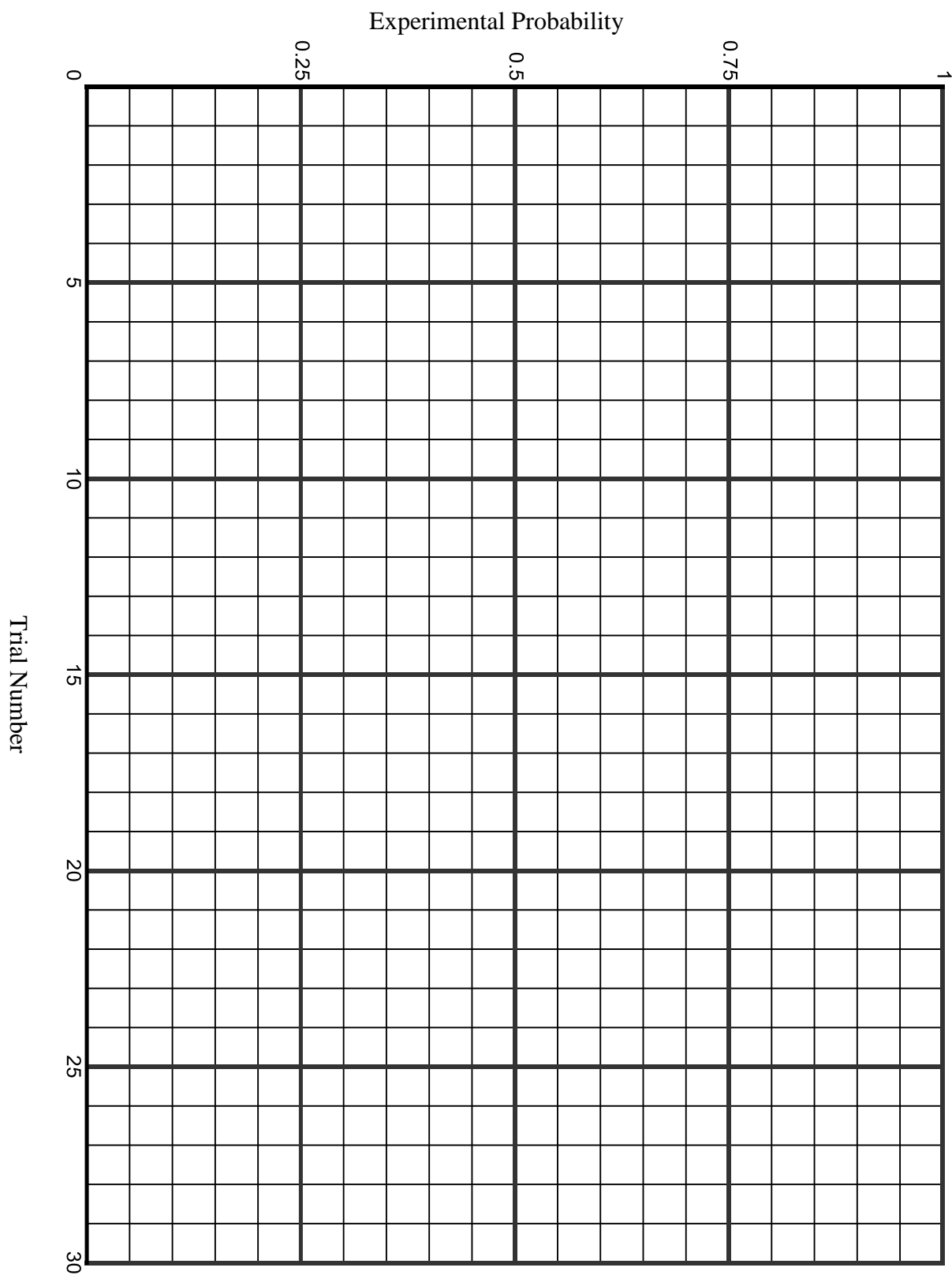
7.1e Homework: Experimental Probabilities

1. Choose an object that has two outcomes when tossed, such as a spoon (face up or face down) or a marshmallow (circular base or side.) Use the techniques from the class activity to find the experimental probability that the object will land on one of the sides.

Plot the data, using the same graph as you used for the class activity. However, plot the outcomes for the homework using a different color, and then label it.

Trial Number	Outcome (B or S)	# of times on Base (running total)	Experimental Probability
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			

Graph your results of # 1 below. Draw lines between the points when you are finished.

Experimental Probability

Trial Number
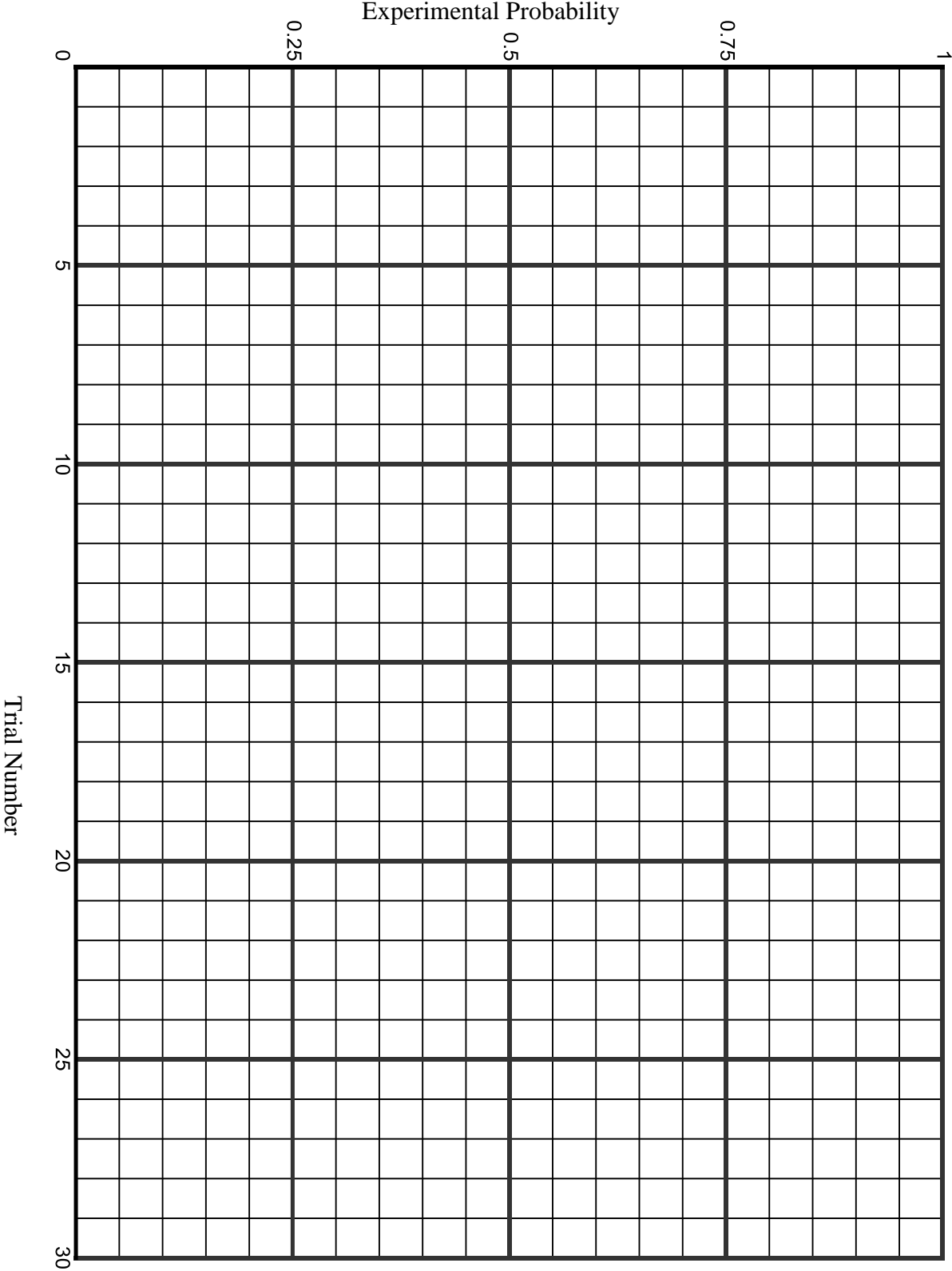
Spiral Review

1. Estimate by rounding to the nearest integer. $-3\frac{1}{3}$, $-\frac{3}{4}$ » ____ , ____ » ____ $\approx$ -3 $\div$ -1 $\approx$ 3

    Is your answer an over estimate or under estimate, explain?   Under estimate

2. David is in a submarine at 200 feet below sea level. Casey is above him in a helicopter at 5,900 feet altitude. How far apart are David and Casey? $5,900 - (-200) = 6,100ft.$

3. Find each difference without a model.
    a. 16 – 29

    b. –2 – (–8)

    c. 5 – (–3)

    d. –90 – 87

4. Given the measures of the following angles, identify the possible angle relationship(s).
    a. $m\angle UTS = 9°$ and $m\angle DCB = 9°$

    b. $m\angle BCD = 71°$ and $m\angle PSV = 109°$

    c. $m\angle PSV = 60°$ and $m\angle GFE = 30°$

5. A baby toy has rings with a radius of 3 inches. What is the circumference of the rings?

7.1f Optional Class Activity: Free Throws or Monty Hall

Activity 1: Free Throws—Will We Win?
You are the coach in the final state basketball championship game; your team is losing by one point. The other team has the ball. You have one of your players foul the person with the ball from the other team. The player from the other team will now shoot two free throws. After the free throws, there will only be enough time to quickly get off a three point shot. The player at the foul line has a free throw percentage of 60%. Your best three point shooter is only a 25% shooter at any three point range.

Your task is to run a simulation to better understand the probability of winning the game.

Using spinners, simulate the situation (spinner can be made out of paper or use internet based sinners, e.g. http://www.mathsisfun.com/data/spinner.php).
*   Spin and record the result of the spin for each of the two free throws.
*   Spin and record the result of the one three point shot.
*   Record if you would win, lose, or tie.
*   Repeat this process 10 times.

1. Based on your simulation, what are your chances for winning?

2. What is the theoretical probability for a win, loss and tie?

Free Throws	3-point Shots	Win, Lose, or Tie

Free Throws

40% miss

60% make

Three point shot

25% make

75% miss
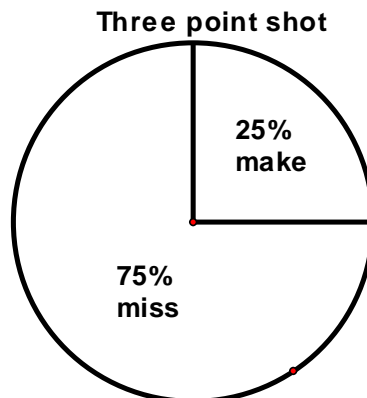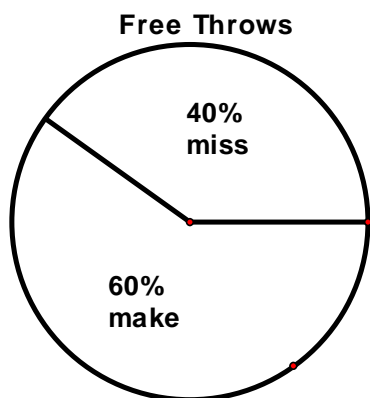
Activity 2: The Monty Hall Question

In the game show Let's Make a Deal, Monty Hall would sometimes show three doors to a contestant. He then informs the contestant that a valuable prize is hidden behind one of the three doors. The contestant would be asked to pick a door. After the contestant chooses a door (without opening), Monty then removes from play one of the doors where the prize is NOT HIDDEN. There are now 2 doors remaining, one of which has the prize. The contestant has already chosen one of these two doors. At this point, Monty gives the contestant the option of switching doors or remaining with his/her original choice.

Which statement below do you think is true:
- The probability of getting the prize is greater by switching doors.
- The probability of getting the prize is greater by not switching doors.
- It doesn't matter.

a. What is your conjecture? Explain.

See: http://betterexplained.com/articles/understanding-the-monty-hall-problem/ for an explanation of the problem and an online simulation. You will want to demonstrate a few rounds of the simulation below so that students understand how to run it. Another option is to use an online simulation as in the link.

b. Run a simulation to see if there is a difference in the probabilities between staying with the same door or switching. Keep track of the results.
- "Monty" rolls a die out of the contestant's view. If the die reads 1 or 2, then the prize (you can use a paper clip or a quarter to represent the prize) is placed behind Door 1. If the die reads 3 or 4, place the prize behind Door 2. If the die reads 5 or 6, place the prize behind door 3.
- The contestant rolls a die. The roll of the die will decide which door the contestant chooses. Use the same numbers as above.
- "Monty" removes from play one of the doors where the prize is NOT HIDDEN. The contestant is asked to remain with the original choice or switch.
- For the sake of consistency, have the contestant REMAIN with the original choice (no switch) for a set number of times. Later, the contestant ALWAYS SWITCHES for an equal number of times.
- "Monty" reveals where the prize is, and the recorder writes down the results in the appropriate column.

c. What are your conclusions?

d. Combine your results with those of the other groups assigned to this problem. What are your conclusions?

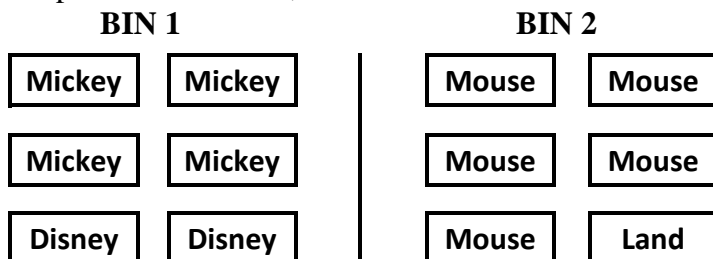e. Explain the results of the game simulation.

7.1f Optional Homework Project: Mickey Match

At the school fundraiser there are two games with prizes. For Game 1 the prize is a Disney decal. For Game 2 there are two different prizes, a t-shirt or a day pass for four to Disneyland. Game 1 costs $1 to pay while game 2 costs $5 to play.

The game is played by picking cards without seeing what is written on the card. The cards are:

Mickey	Mouse		Disney	Land

The cards are placed in two bins, as shown below:

BIN 1		BIN 2	
Mickey	Mickey	Mouse	Mouse
Mickey	Mickey	Mouse	Mouse
Disney	Disney	Mouse	Land

Game 1: To win a Disneyland decal, you pick a card from the left bin. If you pick "Disney" you win a decal. What is the probability of winning a decal? 2/6 = 1/3

Game 2: To play game 2 you must draw one card from Bin 1 and one from Bin 2. Prize options:

> Option 1: If you draw Mickey + Mouse, you win the t-shirt.
> Option 2: If you draw Disney + Land, you win a day pass for four to Disney Land.
> Option 3: If you draw Disney + Mouse or Mickey + Land, you go home with no prize.

Multiple Representations: What is the sample space? Create an organized list, a tree diagram, or a table to see all the possible outcomes.

List:
S = {Mickey + Mouse, Mickey + Land, Disney + Mouse, Disney + Land}

Note that in the sample space the duplicates are not listed. However, students need to be aware that there are duplicates. Prompt the students by asking the number of outcomes for each of the elements in the sample space.

	Mouse	Mouse	Mouse	Mouse	Mouse	Land
Mickey	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Land
Mickey	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Land
Mickey	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Land
Mickey	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Mouse	Mickey + Land
Disney	Disney + mouse	Disney + mouse	Disney + mouse	Disney + mouse	Disney + mouse	Disney + Land
Disney	Disney + mouse	Disney + mouse	Disney + mouse	Disney + mouse	Disney + mouse	Disney + Land

1. Probability Distribution: Find the probability for each outcome and add to your list, table, or tree diagram.

   Mickey + Mouse $\frac{20}{36} \approx 0.56$ Mickey + Land $\frac{4}{36} \approx 0.11$

   Disney + Mouse $\frac{10}{36} \approx 0.28$ Disney + Land $\frac{2}{36} \approx 0.06$

   Note that the sum of the probabilities for each outcome are 1 (or 100%).

2. What is the probability for winning a t-shirt? approximately 0.56, or 56%

3. What is the probability for winning a day pass for four to Disney Land? approximately 0.06, or 6%

7.1g Self-Assessment: Section 7.1

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. Sample problems can be found on the following page.

Skill/Concept	Beginning Understanding	Developing Skill and Understanding	Practical Skill and Understanding	Deep Understanding, Skill Mastery
1. Develop a probability model and use it to find probabilities of events.	I struggle to find probabilities of events.	Given a probability model, I can use it to find probabilities of events.	I can develop a probability model and use it to find probabilities of events.	I can develop a probability model and use it to find probabilities of events. I can show how my model represents the event.
2. Compare probabilities from a model to observed frequencies.	I don't understand the relationship between probabilities from a model and observed frequencies.	I know that probabilities from a model and observed frequencies may be different, but I struggle to explain why.	I can compare probabilities from a model to observed frequencies.	I can compare probabilities from a model to observed frequencies. I can explain possible sources of any discrepancies if applicable.
3. Represent sample spaces for compound events using various methods.	I struggle to represent sample spaces for compound events.	I can represent sample spaces for compound events using one of the following: organized lists, tables **or** tree diagrams.	I can represent sample spaces for compound events using organized lists, tables **and** tree diagrams.	I can represent sample spaces for compound events using organized lists, tables and tree diagrams. I can explain which choice would be best in a given situation.
4. Understand that the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs.	I don't understand how probability is the fraction of desired outcomes in the sample space.	I understand the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs, but I sometimes have trouble applying what I know.	I understand the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs, which can be written as a fraction, decimal, or percent and can use that knowledge in contextual problems.	I understand the probability of a compound event is the fraction of outcomes in the sample space for which the compound event occurs. I can compare this to probability of simple events.

Sample Problems for Section 7.1

1.   For each of the following situations, create a probability model, showing possible outcomes. Then find the probability of the given event.
   a.   Sylvia has a collection of books. She has 30 reference books, 18 nonfiction books, and 64 fiction books. Find P(fiction book).

   b.   The spinner illustrated to the right is spun twice. Find P(white, black).

2.   Don rolled a two on a fair twenty-sided die seven times out of 80 rolls. Would you expect this result? Why or why not?

3.   Represent the sample space for each of the following events. If possible, use various methods for representing the same space.
   a.   Sum from rolling a six-side die twice
   b.   Flipping a quarter four times
   c.   Choosing an outfit out of a plaid, stripped or solid shirt and jeans, khakis or shorts

4.   Crysta puts each letter of her name on a piece of paper. What is the probability that she will draw a C and an A in any order?
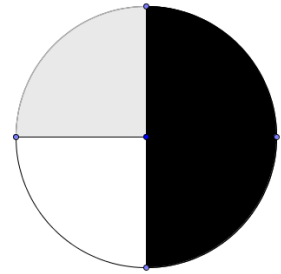
Section 7.2: Use Random Sampling to Draw Inferences about a Population

Section Overview: In this section students will be looking at data from samples and then making inferences from the samples to populations. Students will utilize graphs of data along with measures of center and spread to make comparisons between samples and to make an informal judgment about the variability of the samples. After examining the samples, students will actually make conclusions about the population.

It is important that students think about the randomness of a sample as well as how variations may be distributed within populations. These ideas are quite sophisticated. Activities within this section are designed to surface various ideas about sampling. Teachers, students and parents are strongly encourages (as always) to review the mathematical foundation for a more in-depth examination of the topics within this section.

Concepts and Skills to be Mastered
1. Use random sampling to obtain a sample from a population.
2. Understand that random sampling procedures produce samples that can represent population values.
3. Create appropriate plots of collected data to provide a visual representation of the samples.
4. Compare samples of the same size from a population in order to guage the variation in the samples. Use this variation to form an estimate of range of where a population value might lie.
5. Make predictions about a population, based on the samples.

7.2a Class Activity: Getting Your Opinion

Activity Description: Before starting the surveys, ask students to pretend like you are a rich teacher! Because you are so rich, you are going to take your entire class on a fabulous trip, all expenses paid. Pick two destinations for your class based on what you think might be appealing to very different parts of your class; for example either the Super Bowl or to a Broadway play in New York; the Olympics or Disney Land; an African Safari or a trip to Europe; an after school dance or after school movie, etc. In order to make it even more appealing, discuss with students the pros of each trip. Don't let them campaign for either of the options. Then say: *"The whole class will to go to either _____, or to _____. In order to find out what the class prefers, I am going to take a survey, but I am NOT going to ask everyone. I'm just going to survey a representative sample of the class."* Tell students that they can't say their preference out loud. They must vote silently on a ballot (a piece of paper).

Your teacher will describe an amazing proposal and then ask which one you'd prefer. See activity description **above**. Preparation: cut up sheets of paper into ballots. You'll need about 8 ballots per survey round. Use the ballots to collect opinions using the **four survey samples as described below.** Prepare for survey #4 by making a copy of the class roll, and cut it up so that there is one student's name on each slip of paper. Place the names in a bowl or cup, and stir the names up.

Teacher: look at the sampling methods for each survey and determine if the suggestion is right for your class. You may determine that a different criteria would be better for your students.

<u>Survey 1</u>: Do you think that survey #1 represents the opinions of the class? Why or why not?
Say: "This survey will include all the people in class who have hair that is longer than their shoulders." Hand out survey papers to only those people. Ask students to choose the activity they prefer. Collect and share the results. Students may say that this sample didn't represent the population/class, in fact the sample will usually have a disproportionate number of girls so they are likely correct. Allow students the opportunity to discuss the bias in this survey method and what that will do to the results.

<u>Survey 2</u>: Do you think survey #2 represents the opinions of the class? Why or why not?
Say: "This survey will include people who are wearing earrings." Hand out survey papers to only those people. Ask students to choose the activity they prefer. Collect and share the results. This survey method likely also has the same bias as above (more girls in it than boys). Discuss the bias in this survey and what it will do to the results. At this point, the boys might be getting a little upset, so tell them you'll use a method that includes more boys.

<u>Survey 3</u>: Do you think survey #3 represents the opinions of the class? Why or why not?
Say: "This survey will include people who have shoes with laces." This may result in more male students than female. Discuss with students if the sampling is representative of the class. Discuss biases in this method—for example, people in warmer climates (e.g. Southern Utah, Hawaii, etc.) may wear sandals or girls may be less likely to have shoes with laces. Tell them you are going to make one more effort to get a good representative survey.

<u>Survey 4</u>: Do you think survey #4 represents the opinions of the class? Why or why not?
Say: "This survey will include people who are wearing green." Help students distinguish between "random" and "bias." In the above three surveys, the sampling is not random because the probability of being chosen is not uniform. Samples that are not random, may be biased—as each of the above are. Survey 4 is also not random, but is less likely to be biased (unless green has some significance that day). When we survey, we want the sampling to be random so that it will be unbiased. This is a requisite for the mathematics.

<u>Survey 5</u>: Do you think survey #5 represents the opinions of the class? Why or why not?
Say: "This survey method will include ____ people in the class. In this bowl/cup is the name of every student in the class. I will draw their names randomly from it for the survey." Survey about 10-15 students. Collect and share the results. Ask students to respond to the question prompt above. This method should be the least biased, and most representative of the class opinion. Ask students why this method is the least biased. Discuss why you shouldn't trust surveys that aren't based on random samples.

Which of the five surveys is likely to be most representative of the class opinion? Explain your reasoning.

After the activity you may want to refer to question #3 in the homework, **Inquiring Students Want to Know**, and allow time for students to design their data collection.

Vocabulary that should be discussed during this lesson: population, sample, random sample.

7.2a Homework: Getting Your Opinion

1. You want to determine the most popular brand of shoe among students in your school. Which of the following samples would provide a good representative sample? Explain your choice, and why you didn't choose each of the others.
    a. Ask every tenth student who comes into the school.
    b. Ask ten of the girls on the basketball team.
    c. Ask all the students in your class.
    d. Ask ten of your friends.

2. You are trying to find out who might come to an evening school play performance. Which of the following samples would provide a good representative sample of the community around the school? Explain your choice, and why you didn't choose the others.
    a. Ask fifty people at the local grocery store.
    b. Ask five adults from several randomly selected streets around the school area.
    c. Call random names from the school telephone directory.
    d. Place questionnaires at local stores with a sign asking people to fill them out and drop in a box.

    The least biased method of collecting a sample is (b), because this would represent a random sample of people in the school area. The method in (a) only asks people in grocery stores, and they might have different opinions from people who don't go to grocery stores. The method in (c) might seem really good, but it leaves out people who don't have their phone numbers listed in a phone book. The method in (d) will only sample the people who care enough to fill out the questionnaire, leaving out busy people, or people who don't see the questionnaire.

3. **Inquiring Students Want to Know!** What are you and other students thinking about? Make a list of topics of interest to you and students in your school. For example: What college do students want to attend? Would students prefer starting school early in the morning and getting out early or starting school later in the morning and then staying later in the afternoon? Etc. Choose a question and design a sampling method for collecting data from 10 or more *randomly selected* students. Then collect the data. Write a paragraph describing the results, and why your method is or isn't a representative sample from the population.

Spiral Review

1. Lisa owes her mom $78. Lisa made four payments of $8 to her mom. How much does Lisa now owe her mother? $-78 + 4(8) = -46$ **Lisa owes her mom $46.**

2. Kaylee's Bakin' Kitchen sells fresh bread. The graph to the left shows batches of bread she can make and how much flour it takes. Is it a proportional relationship? If it is, estimate the unit rate? Yes, 7 cups per batch.
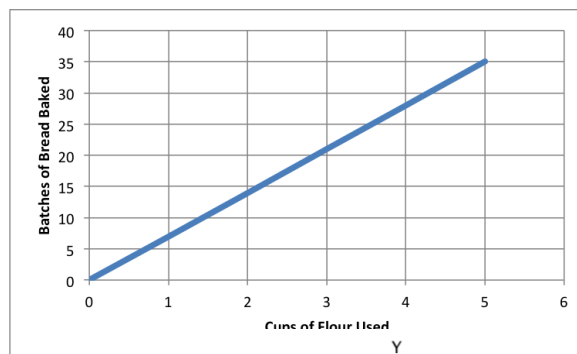
3. What is the scale factor that takes ΔXYZ to ΔABC?

$$\frac{1}{3}$$
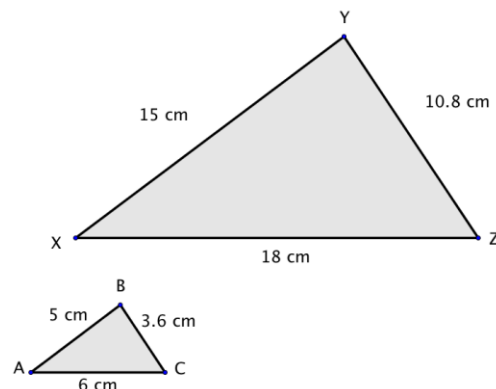
4. Dave is thinking of his favorite number. He tells you that it is one more than three times Emma's favorite number. The sum of their two numbers is 17. What are Dave's and Emma's favorite numbers?

5. Harry's football team loses 13 yards on one play. On the next play, the quarterback throws to a receiver for a gain of 13 yards. What was the change in their position?

7.2b Class Activity: Cool Jelly Beans!

7.2b Class Activity: Cool Beans!

The big election in Jelly Town is coming in November! The Jelly Beans living in Jelly Town (the bag) will cast a vote, either for Limey or for Grapey Bean. Up until now, Limey Bean and Grapey Bean have been tied in the polls. Limey Bean decided that if she wanted to win the election, she needed to do something drastic! So in October she came up with a new campaign slogan "I promise free sunglasses for every Jelly Bean!" Will her new slogan change the way the beans vote? She hired teams of experts to survey the population and answer this question.
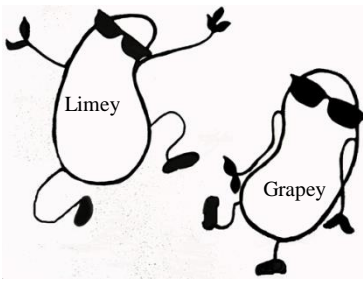
The student groups in class are the experts hired by Limey Bean.

# of green in each sample	0	1	2	3	4	5	6	7	8	9	10
Tally Marks											
Total											

Create a dot plot with your group's data.
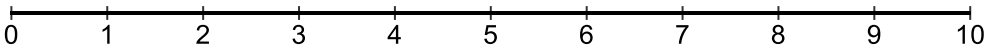
Now that Limey Bean is using the new campaign slogan, there are three possibilities for election results:
a) The new slogan may have made no difference, Limey Bean and Grapey Bean could still be tied.
b) The new slogan may have backfired, so that voters now prefer Grapey Bean.
c) The new slogan may have worked as Limey Bean hoped, so that voters now prefer her over Grapey Bean.

1. Which of the three possibilities (a, b, or c) do your samples support? Justify your answer using your team's survey results. As a hint to students, ask them to find the location on their dot plot representing a tie between Limey and Grapey (at 5, where the number of lime = grape). If the majority of surveys show Limey getting more votes, then it is more likely that she will be the winner of the election. Note also that it is possible for sample results to show Grapey getting more votes some of the time. It is possible (but not likely) to get a sample with more grape beans than lime.

2. Consider your team's survey results. How many of the samples had more votes for Limey Bean? How many samples showed a tie? How many showed more votes for Grapey Bean?
Identify the meaning for specific points on the plots. For example: What does a value of 9 represent? (Limey wins.) What does a value of 4 represent? (Grapey wins.)

3. Based on your samples, find the percent of surveys where Limey Bean had the most votes. Do you think you have enough evidence to declare the winner? Explain why or why not.   Most dot plots should show very few results with Grapey winning. Assuming the sampling was not biased (no one looked at the beans or miscounted), it is justifiable to say: "If the election were held today, it is most likely that Limey Bean will win." You can't say you are certain Limey Bean will win, because of the variability, and because the election is still a month away (it is hypothetically October). You can only say that the chances for Limey Bean winning are pretty good.

4. How variable were the results of your samples? In other words, what was the highest number of green beans recorded from any survey, and what was the lowest number of green beans recorded from any survey? Refer to the numbers found in the table on the first page of the activity. As seen in the table it is highly unlikely to get a sample with only 0, 1, 2, or 3 green beans. Ties happen about 2 out of 20 times or 10% of samples. In large classes with lots of samples, you should see a few 4's and a few 10's. Most survey results will range from 5 through 9.

5. Based on your answer above, is it possible that Limey Bean will lose the election? Is it probable? "Possible" and "Probable" are two different things. Almost anything is possible. It is possible to win the lottery. However, it isn't probable (i.e. likely). So, based on these samples, it is possible Limey will lose the election if it were held today, but it isn't probable.

6. Summarize who you think will win the November election, and why.

7.2b "Cool Beans!" Homework

The campaign in Jelly Bean Town actually began in March with the election in November. The plots below represent surveys taken during the election process in March, August, and October. Each plot shows the results of 20 different surveys**.**
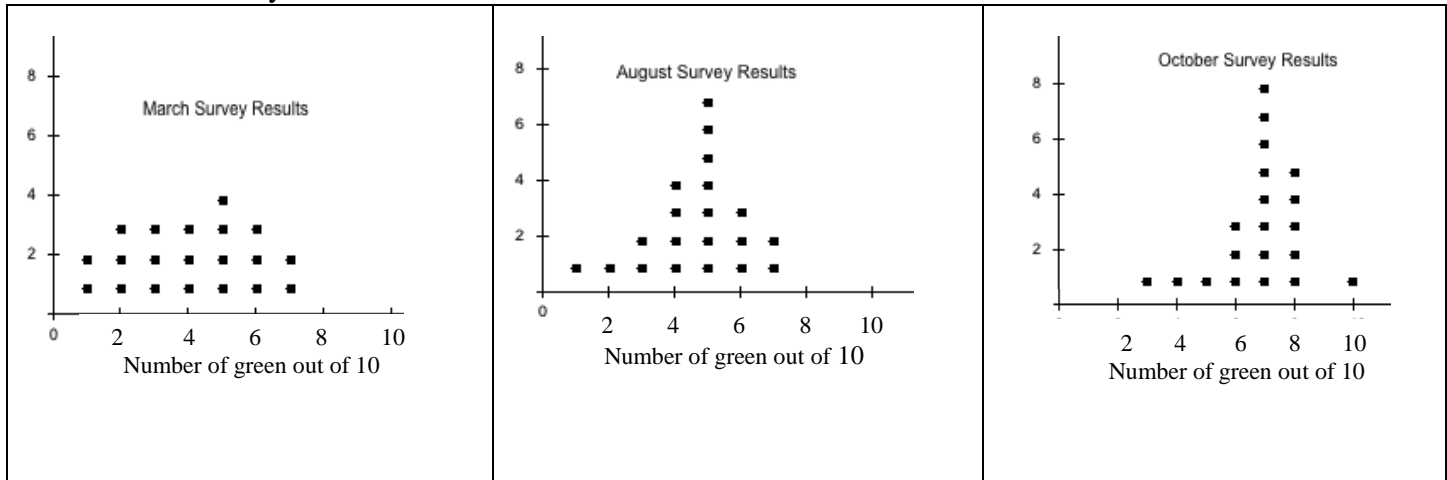

1. In the March surveys, circle the dots on the graph where Grapey Bean and Limey Bean were tied. Based on the graph, who is ahead in the campaign in March? Explain your answer. In March, there are 4 ties between Limey and Grapey (dots which are found over the 5). There are 11 out of 20 dots below the 5, showing Grapey in the lead.

2. In the August results, circle the dots on the graph where Grapey Bean and Limey Bean were tied. Based on the graph, who is ahead in the campaign in August? Explain your answer.

3. In the October results, circle the dots on the graph where Grapey Bean and Limey Bean were tied. Based on the graph, who is ahead in the campaign in October? Explain your answer.

4. Based on the plots, is it possible that Grapey Bean could be ahead in the campaign in October? Explain. It is possible. 2/20 dots show Grapey in the lead. It just isn't very likely that Grapey is actually in the lead.

5. Grapey Bean isn't going to let Limey Bean win the election based on a catchy slogan! In the week before the election Grapey decides to fight back by promising "More Coolness, Less Darkness!" Grapey Bean quickly recruited several expert survey teams to sample the Jelly Bean Town population, in hopes that the new slogan will turn the tide back in Grapey's favor.

After advertising Grapey's new campaign slogan, the three different survey teams gathered data and plotted their results. There is one plot for each survey team.

Was Grapey Bean's slogan successful? Will Grapey win the election now? Use the combined results from the 3 teams' surveys to justify your answer.

Number of green out of 10

Number of green out of 10

Number of green out of 10
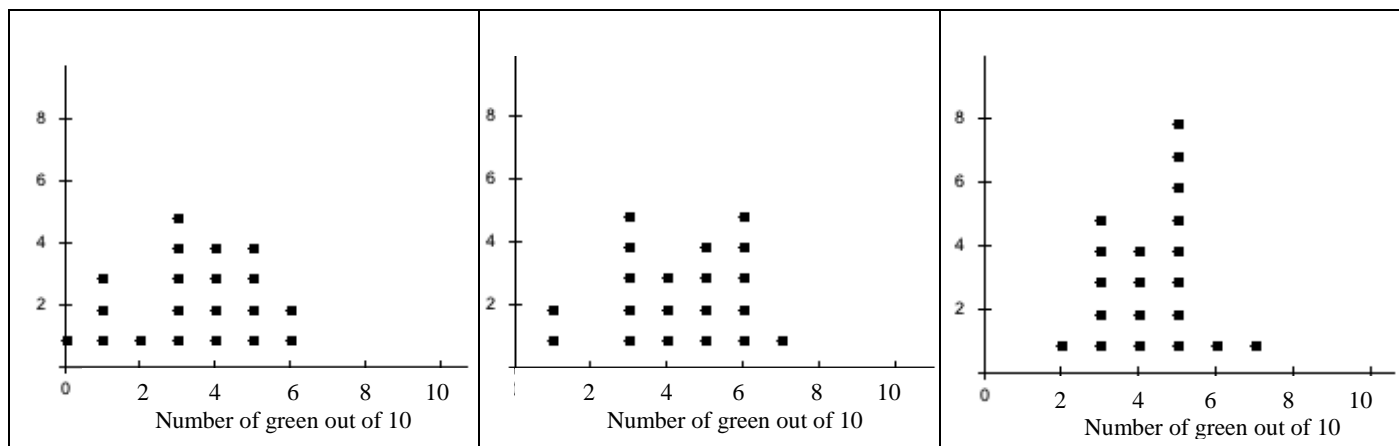
Spiral Review

1. $-1 \times -3 \times -6$ -18

2. Show two ways one might simplify: $2(3 + 4)$

3. Convert 0.37 to a percent.

4. Art's long jump was 7 feet shorter than Bill's. Together they jumped 41 feet. Write and solve an equation to find how far they each jumped? $b + (b - 7) = 41$ Bill jumped 24 ft; Art jumped 17 ft

5. Examine the graph to the right showing how many hours worked versus how much money was made. Explain what the point (2, 20) means in context of the situation and the unit rate.
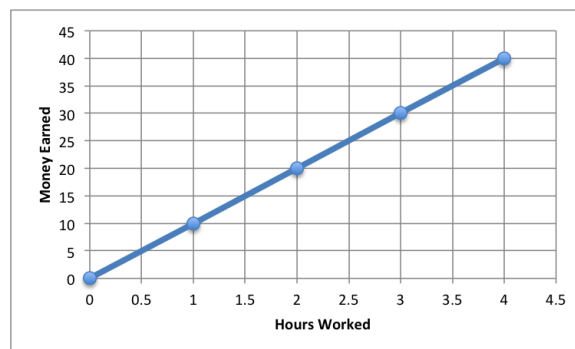
7.2c Class Activity: Critter Sampling (*Optional Activity*) Teacher Notes
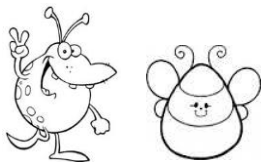
Teacher notes: Scientists often want to study and make estimates about population of plants or animals in a given area. To do this they use random sampling techniques. The data from random samples are used to make conclusions about the populations.
One purpose of this activity is to use samples to draw conclusions about populations and another is to make comparisons between populations. Student groups will be drawing conclusions about their own critter population and then comparing their population to other groups.

Random sampling: Students will be doing a form of random sampling called "cluster sampling". Students will be working in groups of 4 to study their world. The population of each world will be divided into 12 different clusters (see below) and each student will randomly select a cluster to study from their group's world.

Each student will analyze their own sample data, pool their data within the group, and make a conclusion about their world's population.

Prepare bags of critters (the populations) prior to class: Purchase 5 different small food items of similar size to mix together. You'll need one small zip lock bag of the mix per group of 3-4 students (4 preferred). To prepare the critters: use a large bowl to mix the 5 different small food items in equal amounts. Foods might include: stick pretzels, fish crackers, Teddy Grahams, large sized cereal pieces, marshmallows, etc. Choose things that are fairly similar in size. Small items tend to end up on the bottom of the bowl or bag. Don't dump everything into the mixing bowl, hold back a portion of each item to be added later.

Scoop about ¾ cup of well-mixed critters into each bag. Add an additional ¼ cup of *one* of the food items per bag according to the table below, that is, extra marshmallows in one bag, extra Teddy Grahams in another bag, etc. Mix in well. When distributing the bags to the students, don't draw the attention of the students to the differences or similarities between the bags. The idea is that the students will be able to identify which populations are similar and which are different, based on their plots and their tables. Label the bags in some way so that you can remember which ones were the same. Students will be asking this later.

Prepare the Critter Worlds: Copy the Critter World sheet, one per group of students. Students will prepare the world by folding the margins up to create a fence, and taping or stapling the corners so the fence stays up.
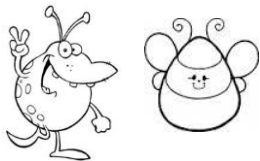
Critter World
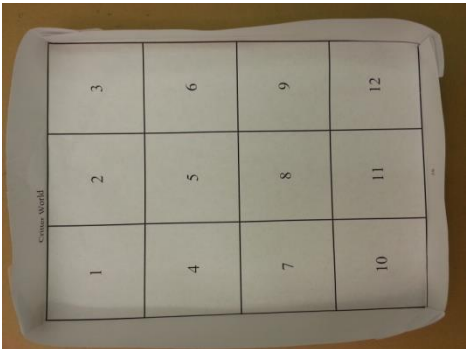
1	2	3
4	5	6
7	8	9
10	11	12

7.2c Class Activity (Optional): Critter Sampling

Your space ship has been orbiting a new flat planet full of life. You are a member of a group of scientists who has been sent to study the flat planet. Your job as an alien biologist is to gather data about the critters, analyze the data, make plots and provide a summary about the area you will study.

The big question is: how is the critter population in your world similar or different from other worlds?
You may choose to do the alphabet activity (homework associated with this activity) in class instead of this activity.
Materials:
Quart sized bags of critters
Sheets of paper, rulers, pencils, tape, scissors, pieces of butcher paper to display graphs

Follow the instructor's directions for setting up the critter worlds. The bags of mixed food items represent different populations of critters. Give your world a name.

1. Open the bag of critters and sprinkle them evenly into the world. Don't use your hands to arrange the critters, just sprinkle them about. Spacing between critters does not have to be exactly equal. If needed, shake the world a bit or stir it with the eraser end of a pencil so that there are critters in every rectangle. Once you are done, hands off! Be careful to not shake or knock the world. (Don't eat the critters until the activity is over!)
    Samples that are not well mixed will give results that can't generalize to the populations. Emphasize the importance of well mixed, random samples.

2. Random Sampling: Each student in the group will randomly select a rectangle to study. To perform a random selection, each group should write the numbers 1-12 on similar sized pieces of paper and then place the numbered slips of paper into a container, mix well, and have each student draw a slip without looking. Some areas of the "world" will not be selected.

3. How should you count critters that are partway between two rectangles? Make a group decision.
    Student methods may vary. They may choose to count a ½ or ¼ critter so that their counts are not whole numbers, they may choose to round to the nearest whole numbers, or they might decide to count a critter as belonging in the rectangle where most of it lies.

4. Count the critters in your rectangle. Record the data in the table.
    Since the critters are not equal in shape or size, don't expect equal numbers of each critter.

Critter Description or sketch	Critter 1:	Critter 2:	Critter 3:	Critter 4:	Critter 5:	
How many?						Total =
Percent of total?						Total = 100%

5. Using the table data, create a graph showing the frequencies of each of the critter types in your sample. (Think: you will be comparing your graphs to others in the group. Why might it be better to use percents rather than counts to make the graph?)

Refer to the Chapter 7 Text for examples of using bar graphs for comparing categorical data and making inferences about populations from the graphs. If comparisons are being made between unequal sized groups, percents give a better basis of comparison. However, if the samples are equal in size, then either counts or percents give comparable graphs. For categorical data such as this, either bar graphs or circle graphs are appropriate. If bar graphs are made, they should be called bar graphs, rather than "histograms" because the data is categorical. Histograms are used for graphs that display a range of numerical values, such as heights or ages. Although pictograms can be used in drawing the graph, with stacks of marshmallows or Teddy Grams, this may generate graphs that don't truly represent the data, unless each pictogram is equal in size.

6. Scientists use data from samples in order to make conclusions about the world. Compare your graph to the graphs made by the other members of your group. Using the data from your samples, come to an agreement on an *estimate* for the total number of each type of critter in your world, and for the percent of each type of critter.

World Population Estimate (count)	Critter 1:	Critter 2:	Critter 3:	Critter 4:	Critter 5:	Total critter estimate =
World Population Estimate (percent)						Does your percent total = 100 %?

7. Describe the method your group used to find the estimates of the world population.

Students may choose to calculate averages for their estimates. However, they might recognize that if sample sizes are different, then averaging the data gives more "weight" to the larger samples, so they might come up with a way to adjust for the different sample sizes. Once they find an estimate, they will need to scale it up to represent the population of the world, such as multiplying their estimate by 12.

8. Within your group, decide how well your samples represented the world population. Explain, using complete sentences, any problems that you might have observed with how the samples may have misrepresented the population.

9. <u>Create a graph</u> of the estimate for the frequencies of the critters in your world. Post this graph in the room. You can eat your critters while all the other groups are posting their graphs.

10. Every world (bag) studied today has at least one other world with a similar population. Look at all the graphs posted by the different groups. Use the graphs to see if you can find the matching sets of worlds. Verify with your instructor to see if you were right!

11. There is variability in between all the samples taken by students. What is "variability" and why did it make matching the worlds challenging? Variability means that every sample will be slightly different. Therefore, instead of finding exact matches between worlds, students have to look for the ones that are most similar.

12. Why is it important to use random sampling and not just *choose* a rectangle to use as your sample? Explain why this would create a problem. Random sampling helps increase the chance that the sample will be an unbiased sample from the population. If people CHOOSE the samples, they might choose based on which sample has the most critters, or which one has the most variety, or which one is in the middle. This creates a bias in the sample, which will also make the population estimate biased.

7.2c Homework: Alphabet Frequency
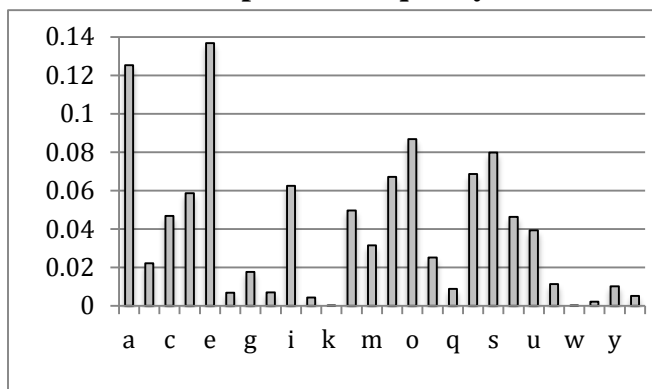
Graph A
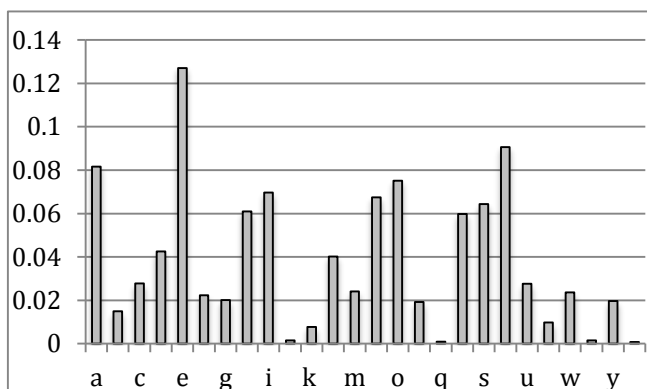
Graph B

The two bar graphs above represent the frequency that letters occur in two languages, one graph represents the English language, the other represents the Spanish language.

1. Which 3 vowels and 3 consonants occurred most frequently in the language represented by Graph A? Write the letters in order from most frequent to least frequent.

2. Which 3 vowels and 3 consonants occurred most frequently in the language represented by Graph B? Write the letters in order from most frequent to least frequent.

3. Use the graph to estimate the frequency of the letter "a" in each language.   Find the difference between the two.

4. The most frequent word in the English language is the word "the". Based on that hint, which graph, A or B, represents the frequency of the letters of the English language? Explain your choice.

Graph B. The "e" is similar in frequency, but the "t" & "h" are much more frequent in Graph B.

5. The bar graph below is a sample of the frequency of letters used in the first two paragraphs of the book *"Artemis Fowl: The Lost Colony"*, by Eoin Colfer, written in English. Since the graph is from a sample, the frequencies will vary a bit from the overall letter frequencies of the English language. Compare the book sample to Graphs A and B. Which graph is the book sample most similar to, A or B?   Explain your choice.

The frequency plot was created on the weblink "Practical Cryptography" at: http://practicalcryptography.com/cryptanalysis/text-characterisation/monogram-bigram-and-trigram-frequency-counts/

7WB 7 - 54
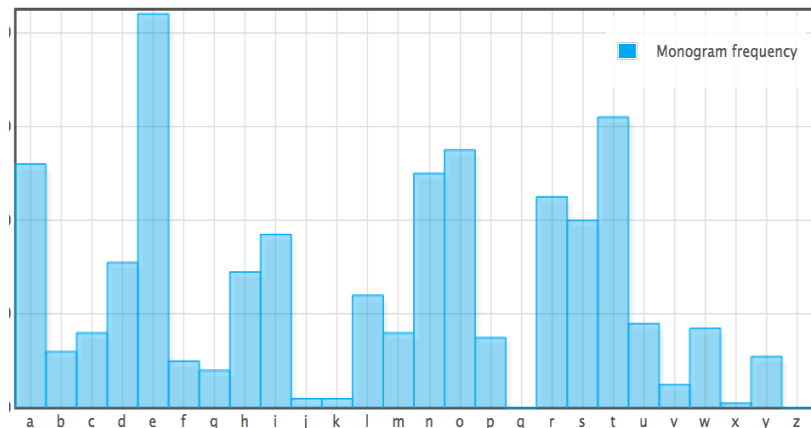
7.2c Homework Extension: Cryptograms

Cryptograms are puzzles where a symbol or letter is substituted for the actual letter. Each of the Artemis Fowl books has a cryptogram at the bottom of the pages of the book, where symbols are substituted for letters, and readers are challenged to solve the hidden message in the cryptogram.

One way people find clues in cryptograms is by looking at letter frequencies. The three most common symbols (in order) from the book "Artemis Fowl: The Lost Colony" are shown below. Which letters do they likely represent?

E, T and A

Use letter substitution to try to solve this famous quote:

YJMIR SKCN MRKCDRMU; MRXS LXIKVX YKNPU.

YJMIR SKCN YKNPU; MRXS LXIKVX JIMZKQU.

YJMIR SKCN JIMZKQU; MRXS LXIKVX RJLZMU.

YJMIR SKCN RJLZMU; MRXS LXIKVX IRJNJIMXN.

YJMIR SKCN IRJNJIMXN; ZM LXIKVXU SKCN PXUMZQS.
                                --WJK-MHX

Hints:
I = c
R = h
S = y
L = b
Y = w
L = b
H = z
J = a
K = o

Spiral Review

1. Determine if the given information will make a unique triangle. Explain why or why not.
   a. Side lengths 10, 10, 19

   b. Angles 51°, 9°, 120°

2. 2. Solve: $3\frac{3}{4} + (-2\frac{1}{2})$. $1\frac{1}{4}$

3. Examine the graph to the right showing ice cream and chocolate syrup needed to make chocolate milkshakes. Is the relationship proportional? If so, write an equation to represent the relationship.

Chocolate Milkshakes
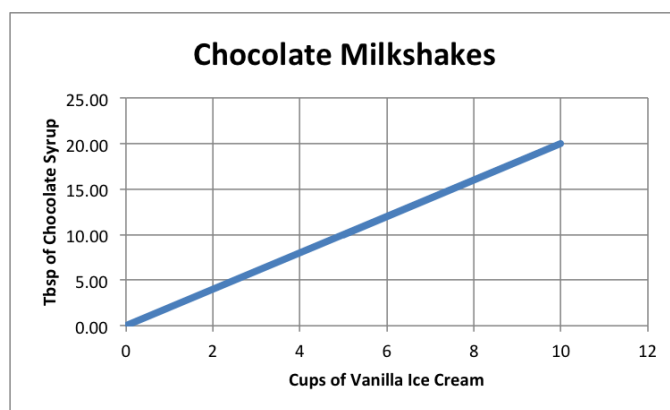
4. Marta is planting a garden as designed to the right. The width of the rectangle is 2 feet. A semicircle is attached to the width of the rectangle. How long should the length of the rectangle be if the total area is 36.56 feet?
   $4x + 3.14(4) = 36.56; x = 6$

5. Two angles are complementary. One is 48 degrees more than twice the other angle. What are the two angles?

7.2d Self-Assessment: Section 7.2

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. Sample problems can be found on the following page.

Skill/Concept	Beginning Understanding	Developing Skill and Understanding	Practical Skill and Understanding	Deep Understanding, Skill Mastery
1. Use random sampling to obtain a sample from a population.	I know what a random sampling is, but I don't know how to use random sampling to obtain a sample from a population.	I can choose which procedure would produce random sampling from a population.	I can use random sampling to obtain a sample from a population. I can explain my procedure for obtaining a random sample.	I can use random sampling to obtain a sample from a population. I can explain why the procedure I used obtains a random sample of a population.
2. Understand that random sampling procedures produce samples that can represent population values.	I don't understand what random sampling is.	I can use random sampling, but I don't understand how that represents the population.	I understand that random sampling procedures produce samples that can represent population values.	I understand and can explain how random sampling procedures produce samples that can represent population values.
3. Create appropriate plots of collected data to provide a visual representation of the samples.	I can't create a plot of collected data.	I can create a plot of collected data, but it doesn't seem to provide a good visual representation of the samples.	I can create a plot of collected data. It is a good visual representation of the samples.	I can create a plot of collected data. I can explain why it is an appropriate plot that provides a visual representation of the samples.
4. Compare samples of the same size from a population in order to gauge the variation in the samples. Use this variation to form an estimate of range of where a population value might lie.	I struggle to compare samples in order to gauge the variation in the samples.	I can compare samples of the same size from a population in order to gauge the variation in the samples, but I struggle to use this variation.	I can compare samples of the same size from a population in order to gauge the variation in the samples, and I can use this variation to form an estimate of range of where a population value might lie.	I can compare samples of the same size from a population in order to gauge the variation in the samples, and I can use this variation to form an estimate of range of where a population value might lie.
5. Make predictions about a population, based on the samples.	I struggle to make predictions about a population, based on the samples.	I can make predictions about a population, based on the samples, but I'm very unsure of my predictions.	I can make predictions about a population, based on the samples.	I can make predictions about a population, based on the samples. I can write a justification for my prediction.

Sample Problems for Section 7.2

1.

    a.   Choose the procedure that would produce a random sampling in the following situation:

        A car insurance company wants to know how many miles people drive each year.
- Ask the teachers at your school how much they drive each year.
- Call every 100th name in the phone book and ask how much they drive each year.
- Ask truck drivers how much they drive each year.

    b.   Describe a procedure that would produce a random sampling in the following situation. Explain why your procedure will produce a random sampling.

        Your school is choosing a new mascot. The principal wants the students' opinions.

2.   Explain how your random samples in question 1 will represent the population.

3.   Belle surveyed her classmates on how many donuts they eat in a month. The following table shows their responses. Make a visual representation of the data.

6	6	6	5	15	19	0	0	0
31	2	3	4	6	6	2	8	8
4	4	4	4	5	5	5	15	15
5	9	2	45	1	1	1	0	5
20	21	25	20	7	7	20	6	7

4. Compare the following two visual representations of how many rings some goblins are wearing. Describe the variation in the samples. Estimate the range of where the population value might lie.
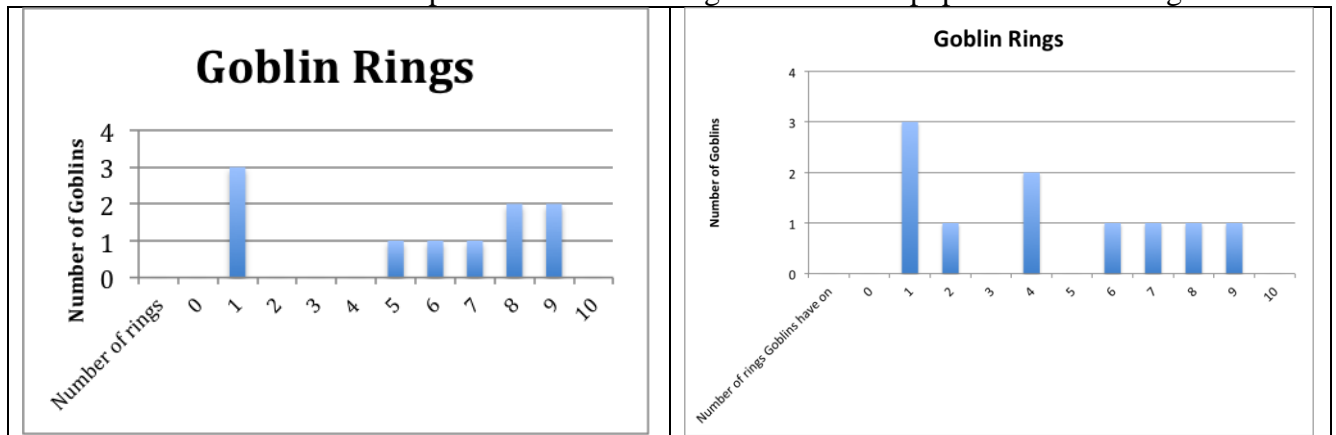
5. Chloe is having a sale on rings in her store. Using the data from the charts in question 4. How many rings should Chloe sell in a set? Explain your reasoning

Section 7.3: Draw Informal Comparative Inferences about Two Populations

Section Overview: In this section students calculate measures of center and spread from data sets, and then use those measures to make comparisons between populations and conclusions about differences between the populations.

Concepts and Skills to be Mastered

1. Make comparisons of data distributions by estimating the center and spread from a visual inspection of data plots.
2. Compare two populations by calculating and comparing numerical measures of center and spread.
3. Calculate the mean absolute deviation (MAD) as a measure of spread of a population. Measure the distance between the centers of two populations of similar variability using the MAD as the unit of measure.

7.3a Class Activity: Viva la Diferencia! (Celebrate the differences!)
Materials needed: One die

This lesson is designed to be an introduction for creating informal comparative inferences about two populations. Students will compare and contrast data gathered from a sample of male students and a sample of female students to determine the sample set which has a larger center and spread and visually identify any outliers. This lesson will rely on prior abilities in calculating measures of center and creating histograms/dot plots.

How do female and male populations compare? With a partner, choose a question below to compare female and male responses. Choose a question that you believe you will find a difference between populations.

a. How many letters are in your first, middle, and last name (total)?
b. How many states can you list in 30 seconds?
c. How many pens or pencils did you bring to class?
d. How many buttons do you have on the clothes you are wearing right now? Include your pants buttons.
e. How many words can you write in 30 seconds that start with the letter "g"?
f. How many minutes does it take you to travel to school in the morning?
g. How many pets do you have?
h. What is the length of your shoe (in centimeters)?
i. How many hours of television do you watch per week?
j. How tall are you (in centimeters)?
k. How long is your hair (in inches)?

Our Question:

_____

_____

Quietly go around the room and record the responses to your question in the table below. When a student asks you their question, you should also ask them your question. Continue until you have at least 10 female responses and 10 male responses.

Boys	Girls

When you have finished gathering your data, go to your teacher to have the die determine how you will display it
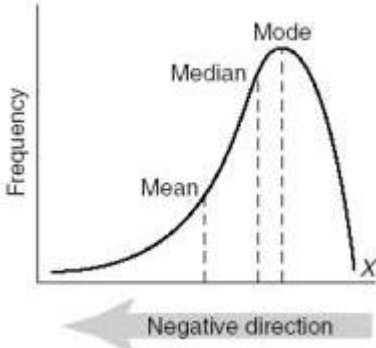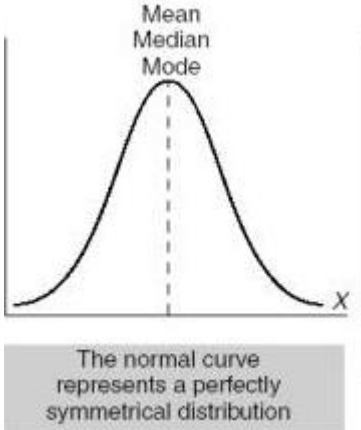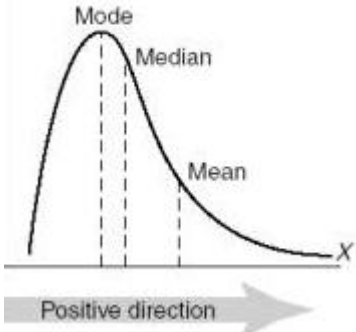
- If the die roll is even, construct a **histogram** to show male and female results.
- If the die roll is odd, construct a **dot plot** to show male and female results.

1. Does the male data or female data have a larger measure of spread? Explain your reasoning.

2. Find the centers of the data for males and females. Which data has the higher center, male or female?

3. Are there any data points that you would consider to be outliers?

4. What conclusions do you draw from the comparison of males and females for your question? Write three to four sentences about your conclusions.
.

Review your work. Prepare to present your data to the class.
Create a poster that has the following:
      1) Your question
      2) A table of your data
      3) Histograms or dot plots of the data (determined from the all-knowing die)

Skewed Left	Normal	Skewed Right or Positive Skew
Extreme values pull the mean to the left of the median. Median is a better measure of center.	Data is symmetrical and mean is at the peak. Mean and median are equally good measures of center.	Extreme values pull the mean to the right of the median. Median is a better measure of center.

7.3a Homework: Review Measures of Center

REVIEW FROM 5th and 6th GRADE:

MODE: The data that occurs with the greatest frequency, or "the most". The mode is an indicator of the shape of a distribution; it is not a measure of center.

MEAN: The mean is a measure of center. To find the mean of a set of data, add all the values together, then divide by the number of values in the data set.

The mean of 18, 6, 0, 22, 5, 19, 7 is calculated by: $\dfrac{18+6+0+22+19+7}{6} = \dfrac{72}{6} = 12$

MEDIAN: The median is a measure of center. To find the median of a set of data, arrange the data in order from least to greatest. If there is an odd number of values, the median will be the middle value. If there are an even number of values, the median will be the midpoint between the values in the middle.

Example 1: Find the median of 33, 35, 10, 19, 7, 0, 0, 6, 7.   Arrange in order: 0, 0, 6, 7, 7, 10, 19, 33, 35
There are 9 data points. The middle (median) value is the 5th one, which is 10.

Example 2: Find the median of 14, 6, 8, 42, 6, 11.   Arrange in order: 6, 6, 8, 11, 14, 42
There are 6 data points. The median is the midpoint between 8 and 11, so the median = 9.5.
The midpoint between two data points can be found by finding the mean of the two points. (8 + 11)/2 = 9.5

Find the mean for the following data sets:

1. 2, 6, 1, 8, 10, 2, 3, 6
> 38/8 = 4.75

2. 24, 14, 8, 9, 6, 5, 18, 10, 16, 22

Find the median for the following data sets:

3. 12, 8, 7, 6, 9, 5, 1, 2, 3
> 6

4. 5, 1.6, 3, 8, 7, 11, 15.5, 18, 20, 11

5. A survey was conducted where respondents gave their favorite summer temperature (in degrees Fahrenheit). The results are as follows: 65, 76, 64, 78, 72, 68, 73, 72, 71, 68, 64, 85, 80, 90. Find the mean temperature from the survey. Round to the nearest degree.
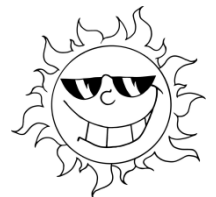
20 males and 20 females were asked to approximate the number of times that they viewed Facebook each day. Histograms for the data are shown below.
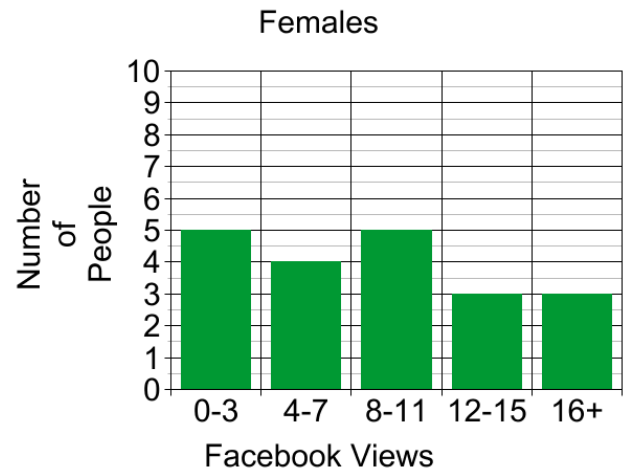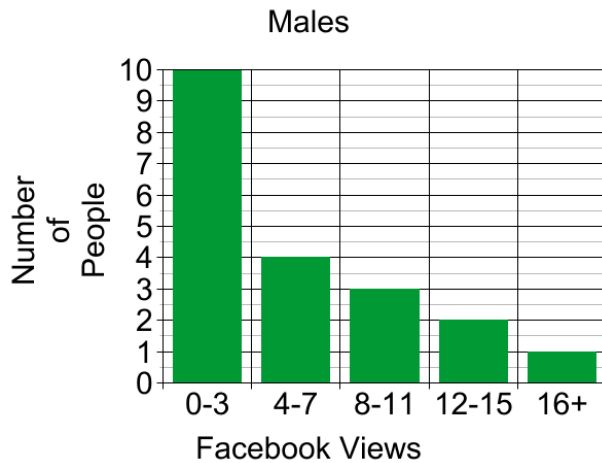
6. Based on those that were surveyed, which group had a greater median, the boys or the girls? Explain your answer. Since there are 20 data points, the median is between the 10th and 11th value.
    The median of the male data is between 3 and 4. The median for the female data is in the 8-11 range. The females' median is higher.

7. Why would mode not be a good measure of center for the female data distribution?

    .

8. Create your own dot plots below that follow these rules:
    Rule #1: Dot Plot #1 must have a larger spread
    Rule #2: Dot Plot #2 must have a greater measure of center

DOT PLOT #1   DOT PLOT #2

Spiral Review

1. Simplify the following expressions.

$$-6(\tfrac{1}{5}a - \tfrac{1}{6}) \qquad\qquad -\tfrac{6}{5}a + 1 \qquad\qquad\qquad \tfrac{1}{3}\left(-7 - \tfrac{1}{6}a\right) \qquad -\tfrac{7}{3} - \tfrac{1}{18}a$$

2. Find each sum, difference, product, or quotient:
   a. $-4 + -7 =$ -11
   b. $3 - 10 =$ -7
   c. $-9(9) =$ -81
   d. $\dfrac{-32}{-8} =$ 4

3. Solve and graph the following inequalities:
   a. $6x - 1 < 17$

   b. $9 \, {}^3 -4x - 7$

4. There were 850 students at Vista Heights Middle School last year. The student population is expected to increase by 20% next year. Draw a model to find what the new population will be. 1,020 students

5. Shawn surveyed his coworkers on how many times they eat out in a month. The following table shows their responses. Make a visual representation of the data.

36	13	19	24	0	12	35	0	30
31	26	3	16	7	27	9	34	27
25	1	35	16	18	17	27	19	6
27	19	11	27	19	26	25	20	23
30	18	26	16	0	0	20	20	18

7.3b Class Activity: The Glorious Mean and Median

Michael Jordan was a professional basketball player in the NBA for 15 years. He is frequently mentioned as the greatest basketball player of all time. He played for the Chicago Bulls team for most of his basketball career. He retired from the NBA in 2003.

The 1997-98 season is one of the years that the Chicago Bulls won the NBA championship. Below is a list of points scored by Chicago Bulls players, from team members who played over 40 games in the season.			The Toronto Raptors basketball team came in last in their division in the 1997-98 season. Below is a list of points scored by team members.		

	Chicago Bulls	**1997/98**		**Toronto Raptors**	**1997/98**
1	Michael Jordan	2357	1	Kevin Willis	1305
2	Toni Kukoc	984	2	Doug Christie	1287
3	Scottie Pippen	841	3	John Wallace	1147
4	Ron Harper	764	4	Chauncey Billups	893
5	Luc Longley	663	5	Charles Oakley	711
6	Scott Burrell	416	6	Dee Brown	658
7	Steve Kerr	376	7	Gary Trent	630
8	Dennis Rodman	375	8	Reggie Slater	625
9	Randy Brown	288	9	Tracy McGrady	451
10	Jud Buechler	198	10	Oliver Miller	401
11	Bill Wennington	167	11	Alvin Williams	324
	TOTAL:	7429	12	John Thomas	151
				TOTAL:	8583

1. Compare the data in the tables without doing any calculations (only using estimates). What interesting features do you see within each data set and between the two data sets? Allow students time to discuss their observations individually before discussing as a class. Students may mention: Jordan has far more points than anyone else, he has about 1300 more points than the next player on his team. He has 1000+ more points than the top player on the Raptors. The Raptors actually scored more total points. Excluding Michael Jordan, the individual Raptors players mostly outscored the Bulls when you compare them side by side, or pair them up player by player.

2. Without calculating the actual values, which team do you think has a higher *points per player* mean? Why do you think so? Students are likely to say the Raptors because of the higher point total. Wait for students to notice that there are more players listed for the Raptors, then ask if this will make a difference in their decision. Students might mention that even with more players, there is still a 1000 point difference in the total points, so the Raptors will still have the higher mean. Challenge the students to estimate the mean for each team before pulling out their calculators for the next question.

3. Calculate the mean and median number of points for each team. (Re-establish the importance of using the correct units of measure in all answers, in this case, "points".)
Bulls' mean = _____ 7429/11 ≈ 675.36 points   Bulls' median = _____ 416 points
Raptors' mean = _____ 8483/12 ≈ 715.25 points.   Raptors' median = _____ 644 points

4. Is the mean or the median a more accurate measure of center for the number of points scored by the Bulls? Explain your choice. Allow students time to think about and discuss this question. The mean and median values are very different for the Bulls, by more than 200 points, so this is a very important question with regards to choosing the right measure of center for this data. When there is an outlying value far away from the rest of the values (like Jordan's), the mean will be pulled away from the center, towards that outlying value. Because of Jordan's scoring, only 4 of the 11 players have score totals higher than the mean, yet the median remains the middle value no matter how many points Jordan makes.

5. What would happen if you replaced Michael Jordan's 2357 points with 10,000 points? Find the new mean and median for the Bull's points per player. Did either value change by much? Explain.
Bulls' Mean = _____1370 points   Bulls' Median = _____416 points
The median stays the same but the mean nearly doubled.
This illustrates why median is a good measure of center if there are outliers. It remains about the same even if there are outlying high or low values. The mean is usually not a good description of the center when there are outlying values.

7.3b Homework: The Glorious Mean and Median

1. Ms. Parrish gave her students a math test and recorded their scores. The following is data for all 16 students in her class: 84, 91, 78, 94, 79, 82, 0, 98, 75, 0, 86, 91, 98, 77, 85, 90. Find the following values:

   a. Mean _____   Median _____   Mode _____

   b. The two scores that are listed as zeros are from students who were absent. Re-calculate the measures of center without the zeros.

   Mean _____   Median _____   Mode _____
   86.29   85.5   91

   c. Explain the effect that the zeros had on the mean, and which values provide the better indication of the center with respect to students' scores.

2. Students tried out for the school play by memorizing a part. The students were rated on how well they performed and how much they were able to memorize. Their ratings were scored on a scale from 0-100. The scores for the 20 students are shown below.

14	79	68	88	84
96	74	94	98	89
97	88	80	94	67
100	98	88	74	88

   a. Sort the data from smallest to largest.

   b. Find the following values:

   Mean _____   Median _____   Mode _____

   c. The first student on the list got a sick stomach during the tryouts and couldn't finish, so only scored a 14. The student was allowed to try again later that day, and now scores a 99.

   What is the new mean score? _____

   How much does the mean score change?

   d. Hamlet is calculating the new mean. Instead of replacing the re-do score of 99, Hamlet adds the re-do score to the end of the list, and then divides the sum by 20. What is result of Hamlet's calculation?
   87.85

   e. Explain why Hamlet's calculation isn't really an average.   What should Hamlet do to fix the calculation? The average should be divided by the total number in the sample. Either Hamlet should have divided by 21, since there are now 21 scores, or Hamlet should have replaced the score of 14 with a 99 and divided by 20. Let the students decide which average is the best representation of the scores for the class.

3. Ten members of the Ceramics Club meet after school to make pottery. A survey was taken to see how far (in city blocks) each member of the club had to travel to get home carrying their heavy pots. The results of the survey are the following distances:

   12, 8, 14, 4, 16, 7, 4, 128, 11, 9

   a. Mean _____   Median _____   Mode _____

   b. Which would be the best measure of center for the data: mean, median, or mode? Explain your

      answer.

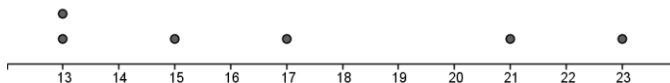   c. Remove the outlier and find the mean of the remaining nine data values.   New Mean _____

      The outlier is the student that travels 128 blocks to school. Eliminating that value gives a mean of 9.44 and a median of 9.22. (Values rounded to the nearest hundredth.)

4. Thomas and Enrique run 2 miles every week and record their times (in minutes). Their data is recorded in the table below:

   Thomas' times

   | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |

   Enrique's times

   | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |

   a. Which runner has data showing the greatest spread? Explain using the plots and comparing data points.

   b. Which runner has the fastest mean time?   The fastest median?
      Thomas's mean is 17 minutes (median is 16.5)   Enrique's mean is also 17 minutes (median is 16). They have the same mean, but slightly different medians.

   c. If you wanted to select one of these runners to represent your class in a running competition, which one would you chose, and why?

5. Caitlin recently took a tour through Mt. Timpanogos Cave. The cave is at an elevation of 6,730 feet and the temperature inside the cave stays a steady 45°F throughout the entire year. Caitlin finds it interesting that the temperature in the cave stays the same year-round. She wonders if the average annual temperature of the air outside of the cave is same or different than the average temperature inside the cave. Caitlin collected data from a nearby community at a similar elevation and found the following typical monthly temperatures for January through December.

   23.6 °, 27°, 34.2°, 42°, 50.3°, 59.1°, 66.4°, 65.6°, 56.6°, 46°, 33°, 24.5°
   Note: Data collected from Kamas, UT (elevation: 6475 feet, and similar latitude as Mt. Timpanogos)

a. Determine if the annual average temperature of the nearby community is the same as the temperature inside the cave. Explain your answer.

b. The Carlsbad Cave system is in the northern Chihuahuan Desert in New Mexico. Steven's family was going on vacation to see the caves. Caitlin told Steven about how cave temperatures seem to be the same as the mean outside temperature. Steven thought he would use Caitlin's information to find out if he would need a coat while he is inside the Carlsbad Cave system. Steven looked up the mean daily temperatures of the nearest large city, which was El Paso, Texas. Although he wasn't able to find the averages he wanted, he found the chart below.

El Paso temperature chart found at: http://www.weather.com/weather/wxclimatology/monthly/graph/USTX0413
Use the chart of the average high and low temperatures of El Paso to find an estimate mean daily temperature of the outside air and help Steven decide if he will need a coat while he is in the caves.
Although the overall average temperature isn't available from the chart, it is reasonable to find the average temperature between the high and low for each month, use all 24 data points and find the overall average, or trace a mid-way line between the high and low values and estimate the average temperature for each month. Any of these methods will provides a fairly accurate estimate.
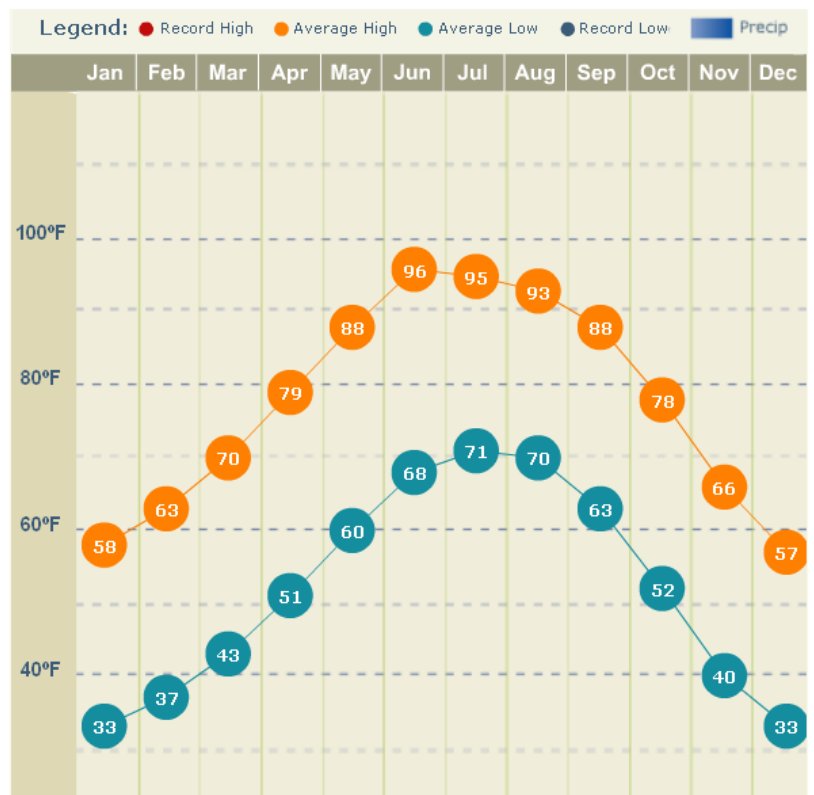
The calculated average is about 65ºF. Steven will probably be okay without a coat.

FYI: The temperature at the deepest point in Carlsbad Caverns is a constant 68°F.

Further information about caves found at the National Park Service website:
http://www.nps.gov/cave/naturescience/weather.htm
"Caves, in general, have fairly stable climate conditions. Once past the entrance area of most caves, the temperature and humidity levels become fairly stable with little variation. This is mostly due to the lack of influence from the outside environment. The temperature in these caves tends to reflect the average annual temperature for the area at that given elevation, though larger cave systems tend to capture some heat rising from the earth's core making them a little warmer than they would be otherwise."

Legend: ● Record High ● Average High ● Average Low ● Record Low ■ Precip

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

Spiral Review

1.   What property is shown?
       a. $19 + 0$ and $0 + 19$ _____identity property of addition___
       b. $9 + 7 + 3$ and $9 + 3 + 7$___commutative_____

2.   Thomas flipped two quarters 80 times. He tails on both quarters 8 times. Would you expect this result? Why or why not?

3.   Willy, Abby, and Maddy are playing golf. Willy ends with a score of -9. Abby's score is -10. Maddy scores +6. What is the difference between the scores of Maddy and Abby?

4.   Beth's golf ball has a circumference of 4.71 in. What is the radius of her golf ball?

Bidziil is examining a scale drawing of the national park near his home. He wants to hike from the park entrance to a hot spring. On the map, the entrance and hot spring are 2.5 inches apart. There is a scale on the map: 1 in = 2.5 mi. How far will he have to hike?

7.3c Class Activity: Got the Point?

1. With a group of 4-5 students, record the number of pens and pencils that each of you have. Write down each of the numbers in the table provided below.

2. Find the mean of your data. What does the mean represent? Answers will vary. The mean represents the average number of pens/pencils for your group.

> **Mean Absolute Deviation** (Review from 6th Grade): The mean absolute deviation (MAD) is a measure of variation in a set of numerical data. It is computed by adding the distances between each data value and the mean, then dividing by the number of data values.

Mean absolute deviation is contained in the 6th grade core. This lesson is meant to revisit the topic and allow students to use that skill in comparing two different populations.

3. Find the mean absolute deviation for the data you collected.

	Number of pens/pencils	Mean All values in this column will be the same	|number – mean| All values should be positive.
Student 1			
Student 2			
Student 3			
Student 4			
Student 5			
		TOTAL:	
		AVERAGE (MAD):	

4. In problem #3 above, you found the mean absolute deviation for your group's data. On the number line below, mark the position of the mean. Put large bracket symbols [ ] above and below the mean at a distance of one mean absolute deviation.

5. Write down the mean and MAD for another group. Like you did above, mark the position of the mean on the number line below. Put large bracket symbols [ ] above and below the mean at a distance of one mean absolute deviation.

6. Is the MAD for your group higher or lower than the other group? What does it mean if a group has a higher MAD?

7.3c Homework: Analyzing Data Using MAD

EXAMPLE for Mean Absolute Deviation (MAD):
The MAD is a measure of the spread of data. The higher the MAD, the more the data is spread out.
The table below shows the 6 fastest birds in the world and their maximum recorded speeds.

Animal	Maximum Recorded Speed (in mph)
Peregrine Falcon	**242**
White-throated Needletail	**105**
Eurasion Hobby	**100**
Frigatebird	**95**
Anna's Hummingbird	**61**
Ostrich	**60**

To find the MAD, first find the mean (average) speed of the birds by adding all of the data and dividing the sum by the number of values.

$$\frac{242+105+100+95+61+60}{6}=\frac{663}{6}=1105$$

Next, find the deviation (distance) from the mean for each bird. Recall that absolute value means you're looking for a "distance" between values and distance is always positive.
Finally, calculate the average of the deviations from the mean, known as the mean absolute deviation, or MAD.

	|speed for each bird – mean|	Deviation from the mean
Peregrine Falcon	|242 – 110.5|	131.5
White-throated Needletail	|105 – 110.5|	5.5
Eurasion Hobby	|100 – 110.5|	10.5
Frigatebird	|95 – 110.5|	15.5
Anna's Hummingbird	|61 – 110.5|	49.5
Ostrich	|60 – 110.5|	50.5
	Mean Absolute Deviation =	43.83

What does the MAD indicate? For this data, the MAD shows that average difference between each bird's speed and the mean is 43.83 mph.

Notice that the Peregrine Falcon's speed is the farthest from the mean. If you use the MAD as a unit of measure, anything that is 3 MAD from the mean is very unusual. The Peregrine Falcon is three MAD's from the mean. $(3 \cdot MAD) = (3 \cdot 43.83) = 131.5$. The Peregrine Falcon is unusually fast, even compared to the other 5 fastest birds in the world!

1. Students in 7th and 9thth grade were asked the number of hours they slept on non-school nights. Data from 20 students in each grade were randomly selected and the histograms for the data are shown below.

7th Grade Students
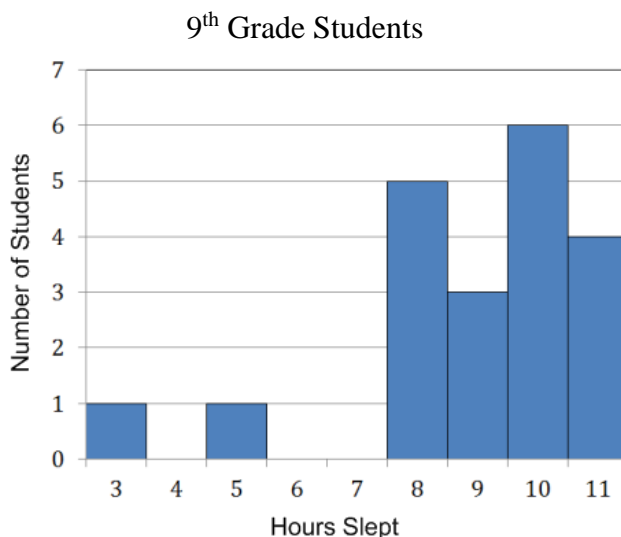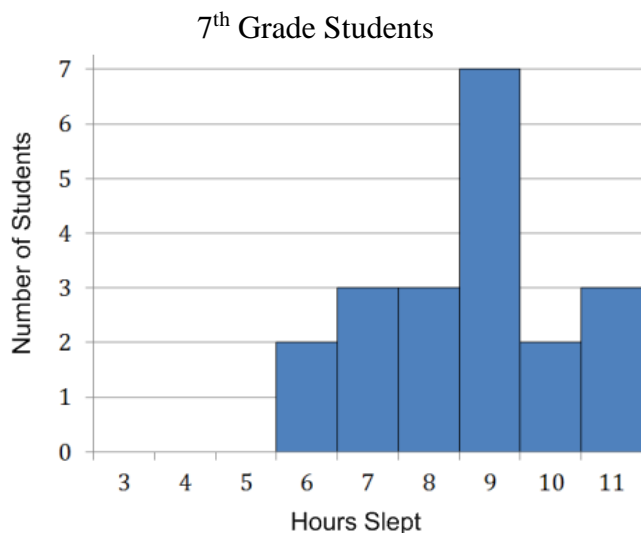
9th Grade Students

1. Find the median of the sleep data for both the 7th grade students and the 9th grade students.
There are 20 students, so the median can be found by looking at the value between the 10th and 11th student.

      7th grade median: _____   9th grade median: _____

2. Without computing, would sleep data for the 7th grade students or the 9th grade students have a larger MAD? Explain your answer. 9th grade has a larger variation, so it would have a larger MAD.

3. Compare the two graphs. Using the graphs and your calculations, write a few sentences about the conclusions that can be made about the amount of sleep that these twenty 7th grade students get compared to the twenty 9th grade students.

4. One statistic used in baseball is how many bases that players steal. This table shows the number of bases stolen by Ken Griffey Jr. and Rickey Henderson each year from 1990 – 2000.

Ken Griffey, Jr. aka "The Kid"

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Stolen Bases	16	18	10	17	11	4	16	15	20	24	6

Rickey Henderson aka "The Man of Steal"

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Stolen Bases	65	58	48	53	22	32	37	45	66	37	36

a. Create dot plots or histograms to provide a visual comparison between the two sets of data.

b. Which player had the highest measure of center for the number of stolen bases? Calculate, and explain your answer.

c. Which player had the greatest spread for the number of stolen bases? Explain your answer by calculating the mean absolute deviation (MAD) for each player.

Ken Griffey, Jr: Mean = 14.3

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	
|number-mean|	1.7	3.7	4.3	3.3	10.3	1.7	0.7	5.7	9.7	8.3	4.7	MAD: ≈ 4.7

Rickey Henderson: Mean = 45.4

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	
|number-mean|												

d. Based on the data, which player would you say has the greater number of stolen bases for their entire career?

Spiral Review

1. Shawn surveyed his coworkers on how many times they eat out in a month. The following table shows their responses. Find the mean, median, and mode of the data.

   Mean: _____

   Median: _____

   Mode: _____

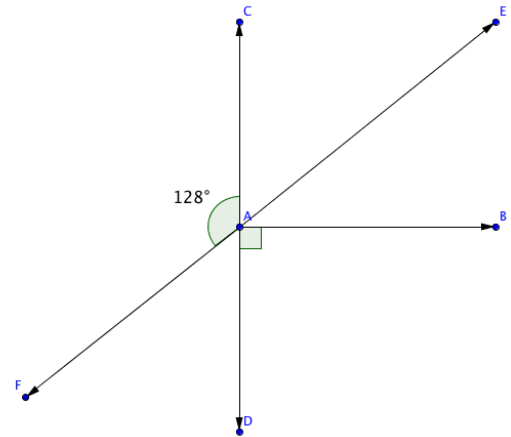2. Find the following quotients:

   a. $\dfrac{3}{5} \div \left(-\dfrac{1}{8}\right)$

   b. $\dfrac{-0.78}{-0.02}$

   c. $\dfrac{-10}{4}$

3. In the diagram to the left, find the missing angles' measures:

angle	measure
ÐBAE	
ÐCAE	
ÐDAF	

4. Ms. Stanford lives in Alaska. When she leaves for work one wintry morning, the temperature is -7° F. By the time she comes home, the temperature has increased 12°. What is the temperature when she comes home?

5. Nellie's bedroom is triangular. She measures the walls as having the following lengths: 10 feet, 10 feet, and 20 feet. How can you tell that she didn't measure correctly?

7.3d Class Activity: NBA Heights
In this activity we will use the MAD to compare the spread of two populations.

Just how much taller are NBA basketball players than students?
You will compare the heights of 25 professional basketball players to the
heights of members of your math class.

Your Height (centimeters)_____

Record your height, and the height of your classmates in the table below.

Basketball Player Heights (centimeters)	
181	204
184	205
186	205
190	206
192	208
194	210
196	210
198	211
199	213
200	214
201	215
202	221
203	

Student Heights (centimeters)		
1] _____	13] _____	25] _____
2] _____	14] _____	26] _____
3] _____	15] _____	27] _____
4] _____	16] _____	28] _____
5] _____	17] _____	29] _____
6] _____	18] _____	30] _____
7] _____	19] _____	31] _____
8] _____	20] _____	32] _____
9] _____	21] _____	33] _____
10] _____	22] _____	34] _____
11] _____	23] _____	35] _____
12] _____	24] _____	36] _____

2. Compare the typical height of students and basketball players:
   a. Calculate the mean of each population. Show all calculations. Round to the nearest centimeter.

   Basketball player's mean height = _____ 202 cm   Students' mean height = _____

   b. How far apart are the mean heights of basketball players and the students, measured in centimeters?

3. Calculate the spread of the student heights :
   a. In the table below, write down the heights of each member in your group. Use the class mean to calculate how much each student in your group varied from the mean.

Student Number	Height (cm)	Deviation From Mean |height – mean|

   b. With the direction of the teacher, record all the other groups responses in the table below.

| Height Deviations from the Class Mean |height – mean|		
1] _____	13] _____	25] _____
2] _____	14] _____	26] _____
3] _____	15] _____	27] _____
4] _____	16] _____	28] _____
5] _____	17] _____	29] _____
6] _____	18] _____	30] _____
7] _____	19] _____	31] _____
8] _____	20] _____	32] _____
9] _____	21] _____	33] _____
10] _____	22] _____	34] _____
11] _____	23] _____	35] _____
12] _____	24] _____	36] _____

c. Calculate the mean absolute deviation (MAD) for the class. Round to the nearest centimeter.

4. The MAD for the heights of the basketball players is 8 cm. Use measure of center, spread and MAD to compare and contrast your class's height to that of the NBA team. Discuss your findings below:

7.3d Homework: NBA Heights

Someone who is more than 3 MAD's shorter (or taller) than the mean height is considered unusual. Use the basketball player height data to answer the following questions about some unusual basketball players.

1. Mark the mean height for the basketball players on the number line (you calculated the mean in the class activity, #2). Measure 1 MAD (8 cm) above and below the mean, and mark each with a "1". Then measure 2 MAD above and below the mean and mark that distance with a "2". Repeat for 3 MAD above and below the mean.
   Mean at 202, 1 MAD from the mean at 194 & 210. 2 MADs from the mean at 186 and 218, 3 MADs from the mean at 178 and 226
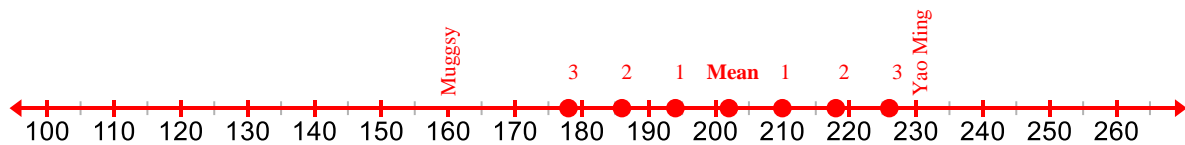
2. Tyrone "Muggsy" Bogues was a professional basketball player from 1987-2001. He was only 5 ft. 3 in tall, which is 160 centimeters. Place a mark on the number line for Muggsy's height, and estimate how many MAD's his height is from the mean.

3. Yao Ming played professional basketball from 2002-2010. He was 7 ft 6 in tall, which is 229 cm. Place a mark on the number line for Yao Ming's height, and estimate how many MAD's his height is from the mean.

4. Whose height was more unusual, compared to the basketball players in this data set, Muggsy's or Yao Ming's?

1. Find the following quotients:

    a. $-\dfrac{3}{7} \div \left(-\dfrac{1}{4}\right)$

    b. $\dfrac{-1.7}{0.1}$

    c. $\dfrac{93}{-5}$

2. Choose the procedure that would produce a random sampling in the following situation:

    > In compiling a brochure about Mathville, the city council wants to know how long people have lived in the city.
    > - Ask everyone in one neighborhood how long he or she has lived there.
    > - Call every 10th name in the phone book and ask how long he or she has lived there.
    > - Ask students at the university how long he or she has lived there.

3. Amie is making cookies for a math party. She has a triangular cookie cutter that is 3 in. on the base and 3.5 in. tall. She rolls her cookie dough into a square with a length of 15 in. About how many cookies will Amie be able to make?

4. Wayne buys a new tie. The tie is 20% off and then he has a coupon for an additional $2 off. If Wayne pays $46, how much was the tie originally?

5. The following list is the names of students in Ms. Jones' kindergarten class. Find the mean, median, and mode for the lengths of their names.

    a.   Mean: _____

    b.   Median: _____

    c.   Mode: _____

Jillian	Justina	Chris	Jodi	Casey
Carl	Nick	Bart	Diego	Doug
Amber	Pat	Kristie	Kaylee	Louise

7.3e Class Activity: MAD about M&M's

Below is data collected for the number of M&Ms in 30 small and 30 large bags of M&Ms. As you can see a small bag contains 1.69 oz. while a large bag contains 3.14, however the actual number of candies varies. The mean and MAD for both small and large bags are also provided. On the next page you will display these data.

M&M Data for Students

SAMPLE #	Small Bag (1.69 oz)	Large Bag (3.14 oz)	SAMPLE #	Small Bag (1.69 oz)	Large Bag (3.14 oz)
1	52	103	16	55	104
2	54	107	17	53	104
3	56	99	18	56	102
4	53	103	19	52	107
5	52	104	20	56	103
6	54	99	21	54	104
7	57	100	22	51	103
8	55	107	23	55	101
9	54	107	24	56	102
10	54	104	25	54	103
11	52	103	26	52	104
12	53	102	27	54	102
13	55	104	28	50	102
14	56	103	29	52	103
15	59	103	30	54	100

	Small Bag	Large Bag
Mean	54	103
MAD	1.47	1.48

In the space below, create a dot plot for the number of M&M's in the small bags of M&M's and create a dot plot for the number of M&M's in the large bags of M&M's.
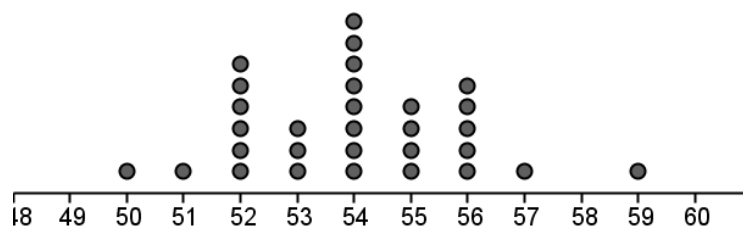
Number of M&M in the small bag.
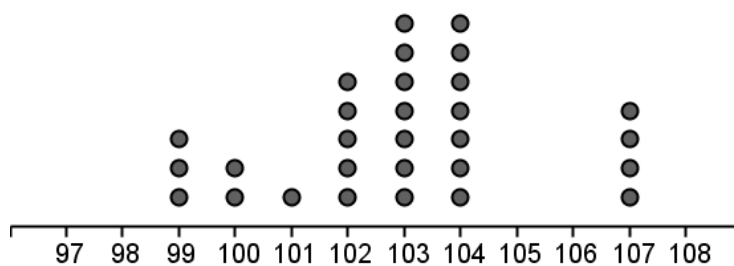
Number of M&M in the large bag.

1. How does the spread for the number of M&M's in a small bag compare to the number of M&M's in a large bag? Explain your answer.

Students may compare the spread based on the dot plots or by comparing their MAD's. Visually students will identify that the 2 sets of data have approximately equal variability and the values for the MAD would confirm that claim. As a class, ask the students what they would think would be a reasonable MAD value. Would a mean absolute deviation of 5 M&M's seem reasonable? 10 M&M's? 1 M&M?

2. On your dot plot below, circle the mean for each data set. What is the difference between the mean for a small bag of M&M's and the mean for a large bag of M&M's?

(see number lines below) 103 – 54 = 49 M&M's; almost an entire small bag of M&Ms
103 and 54 should be identified on the number line.

What is the difference between the centers as a multiple of the MAD value?

3. The mean absolute deviation for both data sets is approximately 1.5 M&M's. Approximate the number of MAD's between 54 and 103.

Have each student reason through the question rather than use a calculator. Have them write down their guess.

Possible student estimation strategies: Students could use the number line in #2 to count the number of MAD's (1.5 units) between 54 and 103. Students may find it easier to mark off every 3 units and then double that number for their answer (see number line below). Other students may use other methods of reasoning to come up with an estimate. Approximations should be around 32 MAD's. ($32\frac{2}{3}$)

Estimate the following:
1. 35/4 ≈ 9
2. 14/3 ≈ 5
3. 35/6 ≈ 6
4. 63/5 ≈ 13
5. 120/8 ≈ 15
6. 1850 / 15 ≈ 123
7. 7/0.5 ≈ 14
8. 654 / 4 ≈ 164
9. 18 / 2.5 ≈ 7
10. 12/1.5 ≈ 8

4. Calculate the number of MAD's between 54 and 103 by dividing the distance between the means by the MAD.

Distance between the means = _____ mean absolute deviations $32\frac{2}{3}$

Explain to the students that the mean of the large bag of M&M's is about 33 MAD's larger than the mean of the small bag.

5. Suppose that a 2.17 oz. bag of Skittles has a MAD of approximately 1.5. There are an average number of 57 Skittles in a bag.

   a. Measure the distance between the means of the 2.17 oz bag of Skittles and the 1.69 oz. bag of M&M's.

        57 – 54 = 3 pieces of candy

   b. Rewrite your answer in part (a) using MAD as the unit of measure.

        _____ pieces   is the same as _____ MAD units
        (Difference in means)
              3   2
        3/1.5 = 2. For every 3 pieces, it is equivalent to 2 MAD.

6. Suppose that a 14 oz. bag of Skittles also had a MAD of approximately 1.5 with a mean of 360 Skittles.

   a. Measure the distance between the means of the 14 oz. bag of Skittles and the 1.69 oz. bag of M&M's.
        240 – 57 = 183 pieces of candy

   b. Rewrite your answer in part (a) using MAD as the unit of measure.
        183/1.5 = 122 MAD

7.3e Homework: MAD About Precipitation

http://www.currentresults.com/Weather/US/average-annual-state-precipitation.php

1. Utah has an average precipitation of 12.2 inches per year, with an MAD estimated of 4.5 inches. Utah is ranked 49th for precipitation out of all the states. (Nevada is 50th.) Precipitation includes both rain and snow.

   a. What does mean absolute deviation (MAD) measure, in terms of precipitation in Utah? The MAD is a measure of how spread out the precipitation amounts are for different locations within the state of Utah. The average absolute distance from the mean is about 4.5 inches of precipitation.

   b. One of the driest cities in Utah is Wendover, getting only 4.1 inches of precipitation per year. How many MAD away from the mean is the precipitation amount for Wendover? 12.2 – 4.1 = 8.1 inches of precipitation. 8.1 inches / 4.5 inches = 1.8 MAD Remind the students that units of measure are a very important part of the answer.

   c. One of the wettest places in Utah is Alta Ski Resort, getting about 54 inches of precipitation per year. How many MAD away from the mean is the precipitation amount for Alta?

2. The state of Hawaii has an average precipitation of 63.7 inches per year, with an MAD estimated of 14 inches. Hawaii is ranked in 1st place for precipitation out of all the states.

   a. One of the driest places in the state of Hawaii is Makena Beach, on the island of Maui. It gets about 17 inches of precipitation per year. How many MAD away from the mean is the precipitation amount for Makena Beach?

   b. One of the wettest places in Hawaii is Hilo, on the island of Hawaii. It gets about 127 inches of rain per year. How many MAD away from the mean is the precipitation amount for Hilo?

3. How much larger is the MAD for precipitation in the state of Hawaii than the MAD for precipitation in Utah? About 3.4 times as large. Estimates are appropriate, so an appropriate answer could also be: between 3 and 4 times as large.

4. Recall that the MAD for precipitation in Utah is estimated at 4.5 inches, and for the state of Hawaii it is estimated at 14 inches. What does that tell you about the range of precipitation values for Utah compared to the range for the state of Hawaii? The range of values for precipitation in Utah is a lot less than for the state of Hawaii, or in other words, the precipitation in Utah is more similar throughout the state than for Hawaii. FYI: the windward sides of the Hawaiian Islands are usually semi-deserts, while the leeward sides are tropical forests.

Spiral Review

1. While living in Mexico City as a foreign exchange student, Ricky kept track of the temperature at noon every day in February. Find the mean, median, and mode temperature in February.
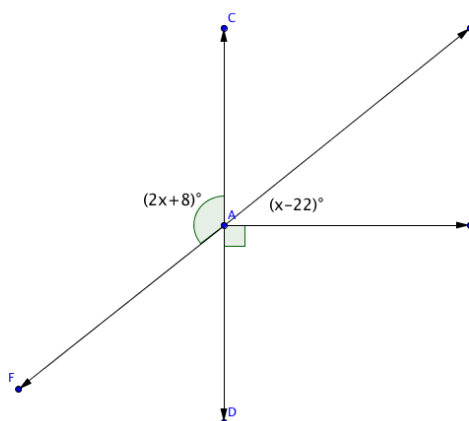
    Mean: _____

    Median: _____

    Mode: _____

55	57	58	52
54	55	57	58
56	59	63	62
65	64	65	64
60	61	62	57
58	54	58	60
59	64	62	63

2. Every morning, Myles picks a random shirt and random pants from his closet. If he has blue, red, brown, and orange shirts and jeans or khakis for pants, what is the probability the Myles will be wearing a brown shirt and jeans?

3. Find the value of x in the diagram to the right:

4. Eden is planting a garden. Her garden plot is $14\frac{1}{2}$ feet long. Strawberry plants should be planted about $1\frac{1}{2}$ feet apart. How many strawberry plants can she fit in one row if she has a $\frac{1}{2}$ foot empty space on each side?

5. Ines is standing on a dock 3 feet above the surface of the lake. She dives down 10 feet below the dock. Then she comes up 7 feet. Where is she now? Write a number sentence showing her movement.

7.3e MAD Olympic Games! -

This lesson is intended to demonstrate that computing MAD values of two data sets allows for the comparison of a single data value in each set to one another.

 For example, suppose that the average amount of ice-cream an American eats per year is 48 pints with a MAD value of 6 pints and the average number of hot dogs eaten per year is 70 with a MAD of 8 hot dogs. Patrick ate 63 pints of ice-cream this year and Monique ate 86 hot dogs this year. Who was more unusual in the amount they ate compared to the mean?   The amount of ice-cream that Patrick ate that year is 2.5 MAD's away from the mean. (63 – 48) / 6 = 2.5 The amount of hot dogs that Monique ate that year is 2 MAD's away from the mean. (86 – 70) / 8 = 2. The conclusion is that Patrick's ice-cream eating is more unusual than Monique's hot dog eating compared to their respective averages.
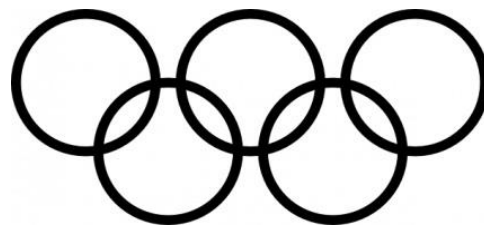
This activity is intended to be a 2-day activity. The first day, the students will be competing in two "Olympic Games", Penny Races and Blind Balance. They will be paired up with another student so that while one is competing, the other is measuring and recording the data.

Instructions for Penny Races:

Supplies needed: Pennies and one timer per student pair (student might use their phones)
Students will be rolling a penny across a tile floor and measuring the distance traveled. The hallway would be a great place for the students to roll their pennies. For ease, have students measure the distance in terms of the number of complete tiles that it was able to travel. No partial tiles allowed. Agree as a class on rounding down to the nearest tile: for example, if the penny goes 10.7 tiles, then the students should record 10 tiles. If you do not have access to a tile floor, you might use the gym. Use masking tape and put a tape mark down for every one foot increment. Label each tape mark based on the foot marking, for example at 4 feet from the starting line, put a piece of tape on the floor that is labeled "4."

Instructions for Blind Balance:

Supplies needed: Timers (again students might use their phone)
Students will be standing crane style: eyes closed, standing on one foot with leg bent in front of body, and arms outstretched. Students will measure the time they are able to maintain the position without dropping their leg, tipping over, or opening their eyes. Students will be in pairs. One students will balance while the other uses the timer and then switch roles and repeat.

Possible Alternate Activities:
If these activities are not feasible, here are some other options:
- Q-tip Javelin Throw – Students toss a Q-tip to see how far it travels. This is similar to the penny races. You may choose to measure distance by the number of tiles or using a tape measure.
- Tongue Tied – Give the students a passage to have the students read, such as the Gettysburg address. Students will record the time it takes them to coherently read through the entire passage. If they cannot be understood, they must restart the passage.
- Paper Clip Puzzle – Each student pair is given 5 paper clips. The students will measure the time it takes to string 5 paper clips together and then take them apart again.

Note that students might feel uncomfortable if they have unusually large or small values. Make note to the students that being an outlier is awesome! It makes the data interesting!

Important!: This is the first day of the 2-day activity. At the end of the first day, collect the student's data. Prior to the next class, compute the mean and MAD for each.

Online Calculator to compute Mean and Mean Absolute Deviation:
http://www.alcula.com/calculators/statistics/mean/
http://www.alcula.com/calculators/statistics/mean-absolute-deviation/

Disclaimer: This activity assumes that the data collected is symmetrical and approximately normal in order to use the mean as a measure of center and mean absolute deviation as a measure of the spread.

7.3f Classwork: MAD Olympic Games!
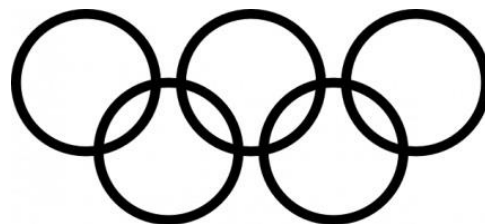
Instructions: As a class, you will be competing in two MAD Olympic events: Penny Races and Blind Balance. You will be working with a partner as you compete.

Penny Races:

With your partner, decide who will be competing first. The goal for this event is to roll a penny as far as you can across the floor.
- Roll the penny across the tile floor.
- Count the number of full tiles the penny traveled from the starting line. Partial tiles don't count!
- Record the data in the table at the bottom of the ★ page by the star for the first student and by the diamond ♦ for the second student...

Blind Balance:

The partner who went second in Penny Races gets to go first in Blind Balance. The goal for this event is to see how long you can stand in the crane position: standing on one leg, with your arms outstretched, and your eyes closed. One partner will balance while the other records time using a timer. Timing ends when your foot touches the ground, you tip over, or if you open your eyes.

- Student #2: Stand on one foot, with your arms outstretched, and your eyes closed
- Student #1: Use a timer to measure how long your partner can stay in the crane position. Record their time in the bottom of the page by the heart ♥ . Round to the nearest second.

Switch roles and repeat. Record Student #1's time by the smiley face. ☺

- -

Once you have finished recording the data, tear out and turn in. Only turn in one for both you and your partner.

PENNY RACES:	NUMBER OF **FULL** TILES:
Student #1:	★
Student #2:	♦

BLIND BALANCE	TIME:
Student #1:	☺
Student #2:	♥

MAD Olympic Games: Day II

Gold Medals in the MAD Olympic Games:

Write down the top records in the two event:

      Penny Races: _____ tiles

      Blind Balance: _____ seconds

"Best of the Best" Title Winner:
1. The MAD Olympic Officials want to give a "Best of the Best" Title. Which winner do you think did the best compared to the rest of the class, the Penny Races winner or the Blind Balance winner?

2. MAD Olympic Officials insist that the "Best of the Best" Title must be given to the player that performed the best in their event compared to the other competitors. The officials have included the mean and mean absolute deviations for each event. Record these values below:

PENNY RACES	BLIND BALANCE
Mean:	Mean:
MAD:	MAD:

Note: Discuss with the students that in order to use the mean as a measure of center and the mean absolute deviation as a measure of spread, the data collected must be approximately normal. For the sake of this activity, have the students assume that the data is approximately normal (which it may or may not be).

3. Ms. Needa Winna suggested that the winner of the title should be the person that has the record that is the farthest from the mean in that event.

      A) Calculate the absolute difference between the Penny Races record and the class mean for the Penny Races Event

            |Penny Races Record – Mean| =

      B) Calculate the absolute difference between the Blind Balance record and the class mean for the Blind Balance Event.

            |Blind Balance Record – Mean| =

Solutions will vary depending on class data. Make sure that the students use correct units when they write their answers. In (A) the units will be in tiles and in (B) the units will be in seconds.

      C) Is this a good method in determining a winner? Why or why not?

This is not a good method. The events have different units of measure, so it is not possible to make a comparison. How would the absolute difference for the Penny Races event be different if it was measured in inches? The following questions will lead the students to convert the absolute differences into MAD units and therefore will allow for comparison.

4. Mr. Hooda Champ suggests that the winner should be whichever event winner has the record that is the greatest number of MAD units away from the mean.

  A) Calculate the number of MAD units away the Penny Race record is from the mean by diving the absolute deviation (#3A) by the Penny Races MAD.

  B) Calculate the number of MAD units away the Blind Balance record is from the mean by diving the absolute deviation (#3A) by the Blind Balance MAD.

  C) Is this a good method in determining a winner? Why or why not?
This is a proper method in determining who is better at their particular event compared to the class. Both records have been converted to MAD units and can therefore be compared. The player that is the largest number of MAD units away from the mean is the most different from the class and is more of an outlier than the other.

5. Who should receive the "Best of the Best" Title? Explain your answer.
Answers will vary depending on the class data. The student who would be given the "Best of the Best" title is the one whose record is the largest number of MAD units away from the mean.
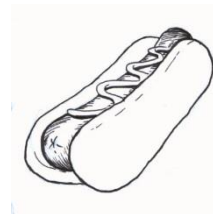
7.3f Homework : MAD Olympic Games

1. Martin participated in a hot dog eating contest. He ate 30 hot dogs in 10 minutes. The average number of hot dogs eaten by the contestants was 12 hot dogs with a MAD of 6 hot dogs.

    a. Martin ate _____ more hot dogs than the average contestant.

    b. Find the number of MAD units Martin was from the mean.
        18/6 = 3

2. According to Pew Internet (2012), teenagers send an average of 60 texts per day. Suppose that the mean absolute deviation is 15 texts. Lily sends about 35 texts per day.

    a. Lily sends _____ fewer texts per day than the average teenager.

    b. Find the number of MAD units Lily is from the mean.

Anya recently got an 85% on her Geography test and a 90% on her Spanish test. She knows that she got a higher grade on her Spanish test, but wonders which test she did better on compared to the class.

3.   How many MAD units away was her Geography test score from the class average?   The Geography test had a mean of 70% with a MAD of 5%.

    |85% − 70%| = 15%   Anya scored 15% higher than the class.
    15% / 5% = 3   Anya's score is 3 MAD units away from the mean.

4. How many MAD units away was her Spanish test score from the class average? The Spanish test had a mean of 80% with a MAD of 5%.

5. Which test did Anya do better on compared to the rest of the class?

Who's more unique??

Joyti Amge – Height: 25 inches	Sultan Kösen – Hand Span: 28 cm
Mean Height: 65 inches	Mean Hand Span: 21 cm
MAD: 3.5 inches	MAD: 2.2 cm

6. Use the data in the table to determine who is more unique, Joyti Amge, the shortest woman in the world, or Sultan Kösen, the man with the largest hand span in the world?

Spiral Review

1.   Find the missing information about the following circle:

        Diameter: _____

        Circumference: 15.7 cm

        Area: _____

2.   Write and solve an inequality for the following problem: The five times the sum of a number and 11 is at most 175.

        $5(x+11) £ 175$

        $x £ 24$

3.   Represent the sample space for each of the following events. If possible, use various methods for representing the same space.
        a.   Sum from rolling a four-side die twice
        b.   Flipping a quarter twice
        c.   Choosing an ice cream sundae from vanilla, chocolate or strawberry ice cream and sprinkles, hot fudge, whip cream or caramel topping

4.   Write and solve an equation for the following problem: Emanuela is arranging her living room. The living room is 10.4 feet wide. Her couch is 7 feet long. How much space should be on each side of the couch for it to be centered along the wall?

5.   The following table shows the amount of dry cereal and water to make hot wheat cereal. Is the relationship of cereal to water proportional? Why or why not?

| Dry Cereal	Water
$\frac{3}{16}$ c	1 c
$\frac{1}{3}$ c	$1\frac{2}{3}$ c
$\frac{2}{3}$ c	$3\frac{1}{4}$ c

7.3g Self-Assessment: Section 7.3

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. Sample problems can be found on the following page.

Skill/Concept	Beginning Understanding	Developing Skill and Understanding	Practical Skill and Understanding	Deep Understanding, Skill Mastery
1. Make comparisons of data distributions by estimating the center and spread from a visual inspection of data plots.	I can make an appropriate plot of data, but I struggle to visually compare two data distributions.	I can compare two data distributions by visually inspecting the data plots.	I can compare two data distributions by estimating the center and spread from a visual inspection of data plots.	I can compare two data distributions by estimating the center and spread from a visual inspection of data plots. I can write an explanation of how they compare.
2. Compare two populations by calculating and comparing numerical measures of center and spread.	I struggle to calculate the center of a population.	I can calculate the center (mean, median, mode) and spread (MAD), but I struggle to use those measures to compare two populations.	I can compare two populations by calculating and comparing the center and spread.	I can compare two populations by calculating and comparing the center and spread. I can write an explanation of how they compare using those measures.
3. Calculate the mean absolute deviation (MAD) as a measure of spread of a population. Measure the distance between the centers of two populations of similar variability using the MAD as the unit of measure.	I struggle to calculate the mean absolute deviation.	I can calculate the mean absolute deviation.	I can calculate the mean absolute deviation. I can measure the distance between the centers of two populations of similar variability using the MAD as the unit of measure.	I can calculate and explain the meaning of the MAD of a population. I can measure the distance between the centers of two populations of similar variability using the MAD as the unit of measure.

Sample Problems for Section 7.3

1. The top ten salaries for sports players in the NBA and NFL are shown in histograms below. Answer the questions that follow.
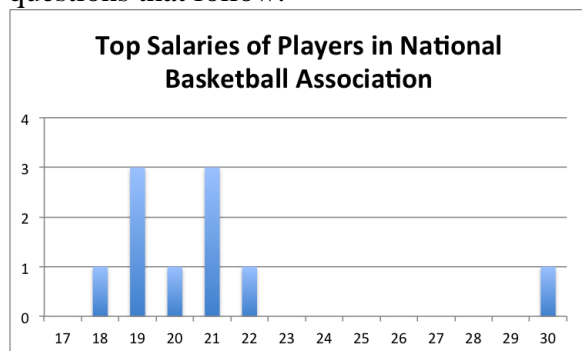
   a.   Which set of data has a higher center?
   b.   Which set of data has a larger spread?
   c.   How do the two data sets compare in similarities and differences?

2. For each data set listed below, calculate the center and spread. The write a comparison of the two data sets. (data from http://www.math.hope.edu/swanson/data/cellphone.txt)

Length of Last Phone Call for Males (in seconds)	Length of Last Phone Call for Females (in seconds)
292	653
360	73
840	10800
60	202
60	58
900	7
60	74
328	75
217	58
1565	168
16	354
58	600
22	1560
98	2220
73	2100
537	56
51	900
49	481
1210	60
15	139
59	80
328	72
8	2820
1	17
3	119

3. Ms. Christensen gave two of her history classes a test. The following table shows the scores from her classes. Find the mean absolute deviation of each. If the MAD is similar for both, find the distance between the centers of the two classes.

5th Period	6th Period
90	72
88	80
100	91
98	90
71	55
83	52
92	76
94	67
87	75
78	80
91	78
99	83
60	78
90	75

Want to find more books like this?

Totally

free kids Books

YAY!

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!