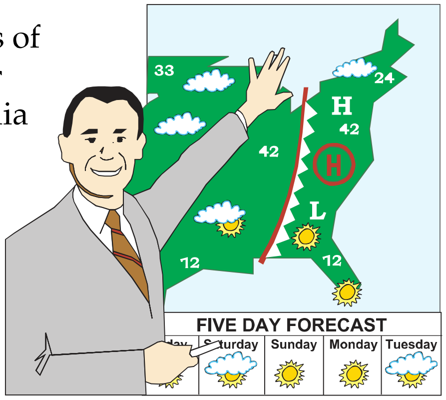
Unit 5: Data Analysis and Probability

Introduction



In analyzing data, we learn to organize sets of statistics in ways that enable us to be better decision makers. The various forms of media which provide us with information frequently use graphs, measures of central tendency (mean, median, and mode), and probability. For example, television weather reporters can predict for a week at a time the probability of rain for a given geographic area.

Therefore, in order to keep up with current events, a person needs to be able to formulate hypotheses, collect and interpret data, and draw conclusions based on statistics, tables, graphs, and charts. Furthermore, we need to be able to recognize ways in which statistics can be misleading. Clever statisticians can devise graphs, charts, etc., that are deceptive. With a good knowledge of data analysis, people can develop skills for interpreting and evaluating statistical presentations.

Lesson One Purpose

- Describe, analyze, and generalize relationships, patterns, and functions using words, symbols, variables, tables, and graphs. (MA.D.1.4.1)
- Determine the impact when changing parameters of given functions. (MA.D.1.4.2)
- Represent real-world problem situations using finite graphs. (MA.D.2.4.1)
- Use systems of equations and inequalities to solve real-world problems graphically and algebraically. (MA.D.2.4.2)

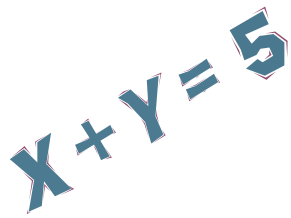
- Interpret data that has been collected, organized, and displayed in charts, tables, and plots. (MA.E.1.4.1)
- Analyze real-world data and make predictions of larger populations by using the sample population data and using appropriate technology, including calculators and computers. (MA.E.1.4.3)

Equations in Two Variables

In this unit, we will study **equations** with two **variables**. We will limit our study to equations such as these:

$2x + y = 10$; $y = 3x + 5$; and $y = 4x$.

To begin, let's see how we can solve an equation like



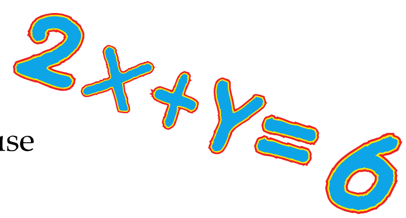
$x + y = 5$.

You will find that, unlike the earlier equations in Unit 3, this type of equation has an **infinite** number (*no limit* to the number) of **solutions**. A *solution of an equation with two variables* is an **ordered pair** of numbers that make the equation true.

Suppose that we replace the variable x with the **value** of 0. Obviously, the variable y would have to be replaced with the number 5, because

$0 + 5 = 5$.

So if $x = 0$, then $y = 5$, and we use the *ordered pair* (0, 5) to denote the solution.



Here are two other solutions for $x + y = 5$.

- If we let $x = 1$, then y would be 4 because

  $1 + 4 = 5$.

Hence, (1, 4) is a solution.

- If we let $x = 2$, then y would be 3, because

  $2 + 3 = 5$.

Therefore, (2, 3) is also a solution.

Many ordered pairs work for the equation $x + y = 5$.

Below is a **table** *of values* with several examples of ordered pairs that are solutions to the equation. A *table* is an orderly display of numerical information in rows and columns.

$x + y = 5$	
0	5
1	4
2	3
3	2
1	4
5	0
$\frac{1}{2}$	$4\frac{1}{2}$
0.25	4.75
⋮	⋮

Example 1: Find five solutions for the equation $y = x - 2$.

Solution: Begin by picking *values* for x. Any number will do.

Equation	x	Substitute for x	Solve for y	Solution
$y = x - 2$	0	$y = \mathbf{0} - 2$	-2	(0, -2)
	1	$y = \mathbf{1} - 2$	-1	(1, -1)
	2	$y = \mathbf{2} - 2$	0	(2, 0)
	-1	$y = \mathbf{-1} - 2$	-3	(-1, -3)
	-5	$y = \mathbf{-5} - 2$	-7	(-5, -7)

Remember: In an ordered pair, the value of x is listed first, and the value for y is listed second—even if the variable y is written first in the equation.
Example: In the equation $y = x - 2$, the solution is (0, -2).

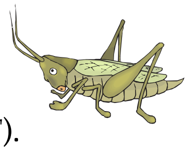
$$-2 = 0 - 2$$
$$-2 = -2$$

$x \quad y$

Example 2: Find three solutions for the equation $2x + y = 6$.

Solution: Pick three values for x. Again, any number will do.

Equation	*x*	Substitute for *x*	Solve for *y*	Solution
$2x + y = 6$	0	$2(\mathbf{0}) + y = 6$	6	(0, 6)
	1	$2(\mathbf{1}) + y = 6$	4	(1, 4)
	-1	$2(\mathbf{-1}) + y = 6$	8	(-1, 8)



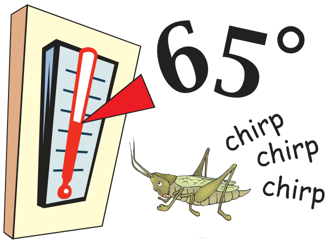
Example 3: $T = \frac{1}{4}C + 40$ shows the relationship between the number of times a cricket chirps per minute (C) and the temperature in degrees Fahrenheit (T). Solve this equation by finding three solutions.

Solution: Pick any values for x. Numbers divisible by four would be the easiest to work with.

Equation	*T*	Substitute for *C*	Solve for *T*	Solution
$T = \frac{1}{4}C + 40$	4	$T = \frac{1}{4}(4) + 40$	41	(4, 41)
	16	$T = \frac{1}{4}(16) + 40$	44	(16, 44)
	40	$T = \frac{1}{4}(40) + 40$	50	(40, 50)

We just found that when a cricket chirps 40 times, then it is 50 degrees Fahrenheit.

Example 4: Using the information in example 3 on the previous page, is it true that when the temperature is 65 degrees Fahrenheit that a cricket will chirp 60 times?



Solution: We basically need to check to see if the ordered pair (60, 65) is a solution for the equation $T = \frac{1}{4}C + 40$.

$$T = \frac{1}{4}C + 40$$

Substitute $65 = \frac{1}{4}(60) + 40$

$$65 = 15 + 40$$

$$65 = 55$$ This is *not* a true statement.

We now know that when the temperature is 65 degrees, a cricket will *not* chirp 60 times. Can you see that the cricket will only chirp 60 times when the temperature is 55 degrees Fahrenheit?

$$T = \frac{1}{4}C + 40$$

Substitute $55 = \frac{1}{4}(60) + 40$

$$55 = 15 + 40$$

$$55 = 55$$ This is a true statement.

Example 5: Which ordered pair (-8, -5) or (12, -10) is a solution for the equation $y = \frac{3}{4}x + 1$?

Solution: Substitute the values for x and y into the equation:

(-8, -5)

$y = \frac{3}{4}x + 1$

Substitute $-5 = \frac{3}{4}(-8) + 1$

$-5 = -6 + 1$

$-5 = -5$ This is true, so (-8, -5) is a solution.

(12, -10)

$y = \frac{3}{4}x + 1$

Substitute $-10 = \frac{3}{4}(12) + 1$

$-10 = 9 + 1$

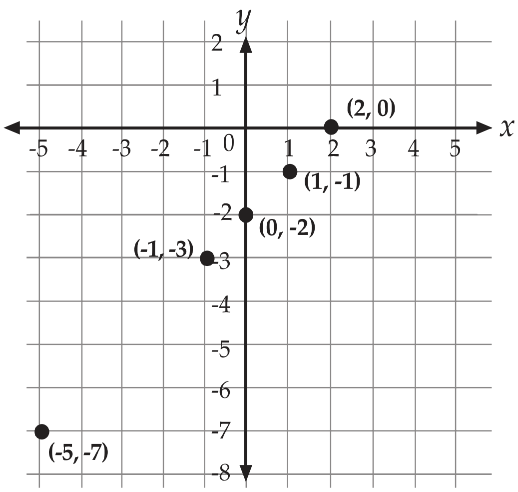
$-10 = 10$ This is false, so (12, -10) is not a solution.

Graphing Linear Equations

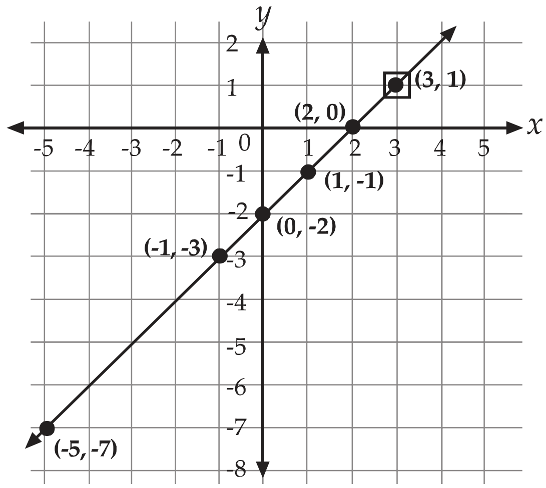
If you refer to example 1 in the last section, you will see that we found five solutions for the equation $y = x - 2$. Here is a *table of values* with the summary of the results.

x	y	
0	-2	(0, -2)
1	-1	(1, -1)
2	0	(2, 0)
-1	-3	(-1, -3)
-5	-7	(-5, -7)

The **graph of an equation** with two variables is all the **points** whose **coordinates** are solutions of the equation. The *coordinates* correspond to points on a **graph**. Let's graph these *points* and see what we get.



Carefully draw the line that connects these points.



The graph of the equation $y = x - 2$ is the **line (⟷)** drawn on the **coordinate plane**. The **coordinate grid** itself contains the ***x*-axis**—the horizontal (⟷) axis and ***y*-axis**—the vertical (↕) axis. The *line* drawn is endless in length and the **plane** is a flat surface with no boundaries.

Notice that the point (3, 1) lies on our line, but it was *not* one of our original points. Let's see if it is a solution. **Substitute** or *replace* the variables x and y with (3, 1).

$y = x - 2$	*substitute* 3 for x and 1 for y
$1 = 3 - 2$	
$1 = 1$	This is true, so (3, 1) is a solution of the equation.

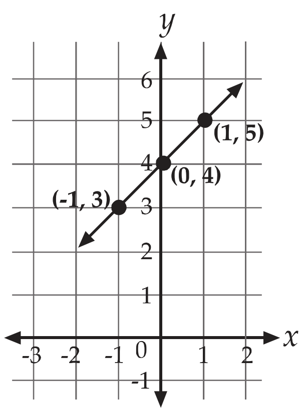
It turns out that any point on the line is a solution of the equation. We cannot write all the solutions to an equation because x can be anything, but we can draw a *picture* of the solutions using the *coordinate plane* and a line. The equations that we have been working with in the last section are called **linear equations** because their graphs are always *straight lines*.

Example 1: Graph $y = x + 4$.

Solution: It only takes two points to determine a line, so we only need two solutions of the equation. Generally, we find three solutions. (If we make an arithmetic mistake, it will be obvious because we will not get a line.) Choose three small values for x. Remember x can be any number that you wish.

Equation	x	Substitute for x	Solve for y	Solution
$y = x + 4$	0	$y = 0 + 4$	4	(0, 4)
	1	$y = 1 + 4$	5	(1, 5)
	-1	$y = -1 + 4$	3	(-1, 3)

Plot these points and draw the line.



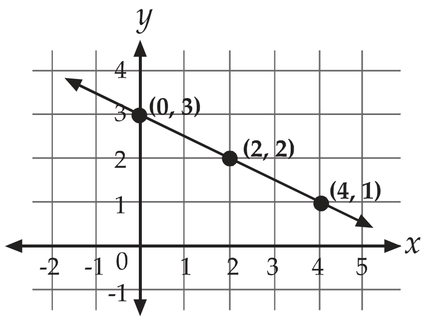
Remember: Any point on the line is a solution.

Example 2: Graph $y = -\frac{1}{2}x + 3$.

Solution: Choose values for x that are divisible by 2.

Equation	*x*	Substitute for *x*	Solve for *y*	Solution
$y = -\frac{1}{2}x + 3$	0	$y = -\frac{1}{2}(0) + 3$	3	(0, 3)
	2	$y = -\frac{1}{2}(2) + 3$	2	(2, 2)
	4	$y = -\frac{1}{2}(4) + 3$	1	(4, 1)

Plot these points and draw the line.



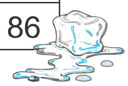
Remember: Any point on the line is a solution.

Interpreting Data and Lines of Best Fit

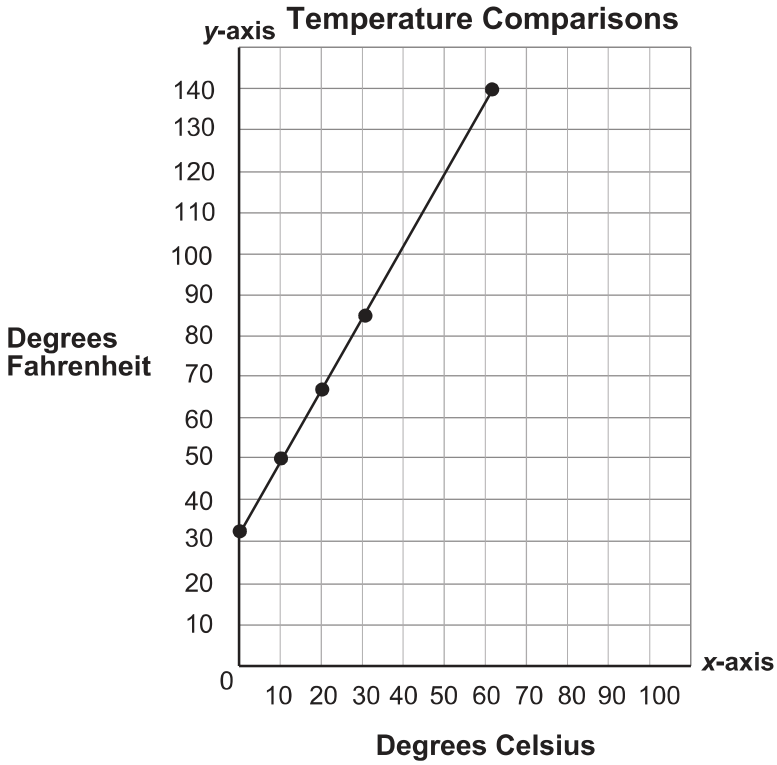
In many situations, the relationship between two variables may be of interest. The following table shows the relationship between Fahrenheit and Celsius temperatures. For instance, a temperature of 0 degrees Celsius corresponds to a temperature of 32 degrees Fahrenheit. This is when water freezes.

Temperature Comparisons

C	0	10	20	30
F	32	50	68	86



Let's graph these ordered pairs. We will graph the Celsius temperatures on the *horizontal* (↔) *x-axis* and the corresponding Fahrenheit temperatures on the *vertical* (↕) *y-axis*. We will carefully draw a line that contains these points.

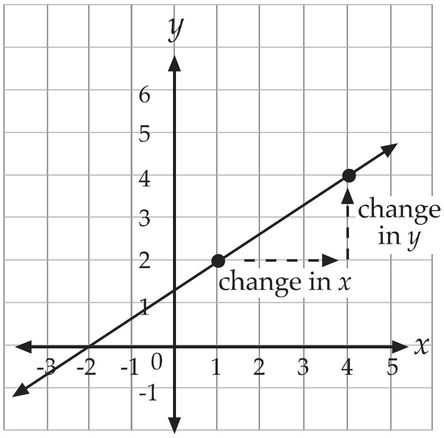


Note that when the Celsius temperature is 60 degrees, the corresponding Fahrenheit temperature is close to 140 degrees, which happens to be the correct answer. When the temperature is 80 degrees Fahrenheit, what would the corresponding Celsius temperature be? We hope you guessed that the corresponding Celsius temperature is around 27 degrees.

The graph on the previous page is a visual representation of **data**. Such a graph is called a **line graph**. The *line graph* shows change over time. The *data* used is information in the form of numbers gathered for statistical purposes. The line graph, which is one type of **data display**, allows a person to see patterns that may not be obvious from just the equation. A graph is a powerful visual tool for representing data.

Slope

We all know that some mountains are steeper than others. Lines in a *coordinate plane* also have steepness. In math, the steepness of a line is called its **slope**. The *vertical* (↕) change is called the *change in y* and the *horizontal* (↔) change is called the *change in x*.



Two Ways to Find the Slope of a Line

The slope or steepness of a line can be found using two points from the line. We will explore two definitions of slope and ways to find the slope of a line.

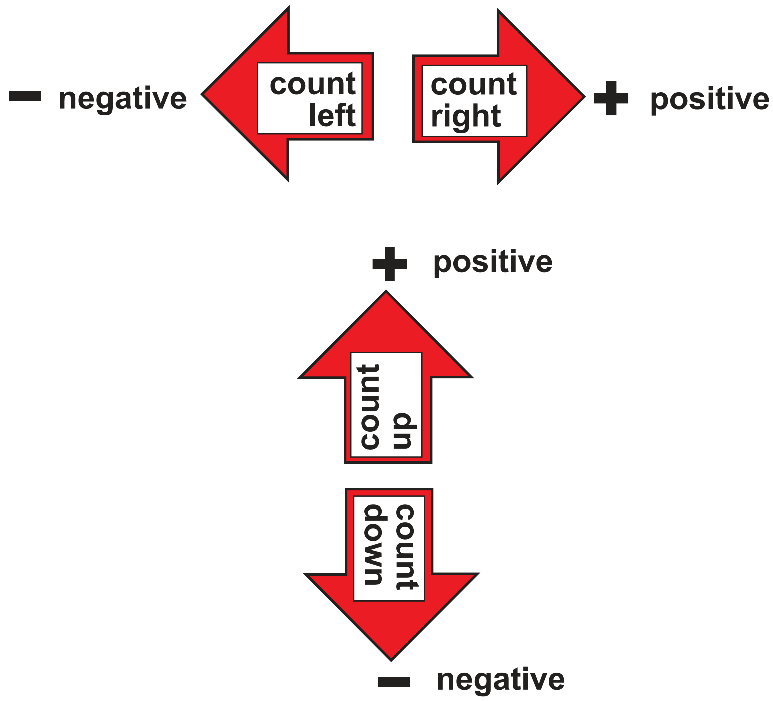
First, *slope* is defined to be the following **ratio**:

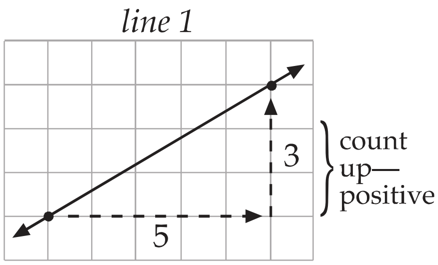
$$\text{slope} = \frac{\text{change in } y}{\text{change in } x}$$

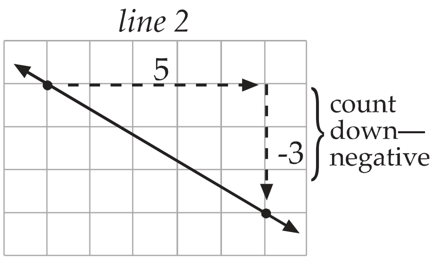
A *ratio* is the quotient of two numbers used to compare quantities.

Example 1: Find the *slope* of the following six lines.

Solution: Find the *change in y* and the *change in x* by counting units. If you count to the right, that is a positive direction. If you count to the left, that is a negative direction. If you count up, that is a positive direction. If you count down, that is a negative direction. Pick any two points on the line, and travel from one point to the other.

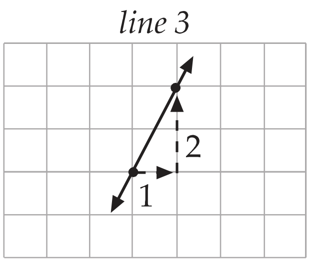


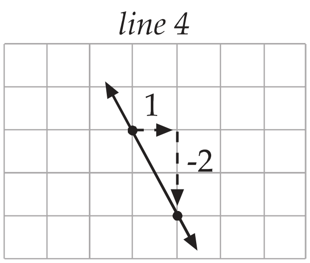




a. slope = $\frac{\text{change in } y}{\text{change in } x} = \frac{3}{5}$

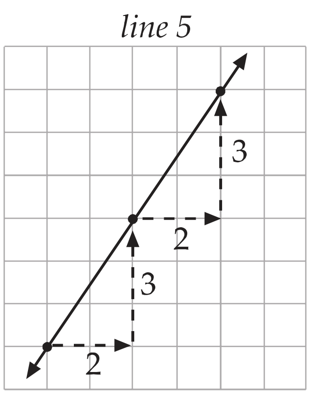
b. slope = $\frac{\text{change in } y}{\text{change in } x} = -\frac{3}{5}$

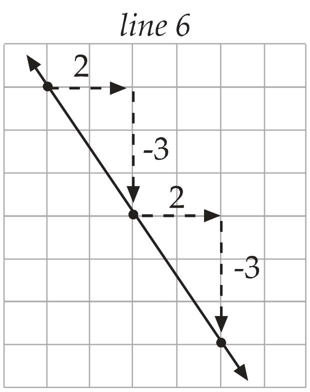




c. slope = $\frac{2}{1}$ or 2

d. slope = $-\frac{2}{1}$ or -2





e. slope = $\frac{3}{2}$

f. slope = $-\frac{3}{2}$

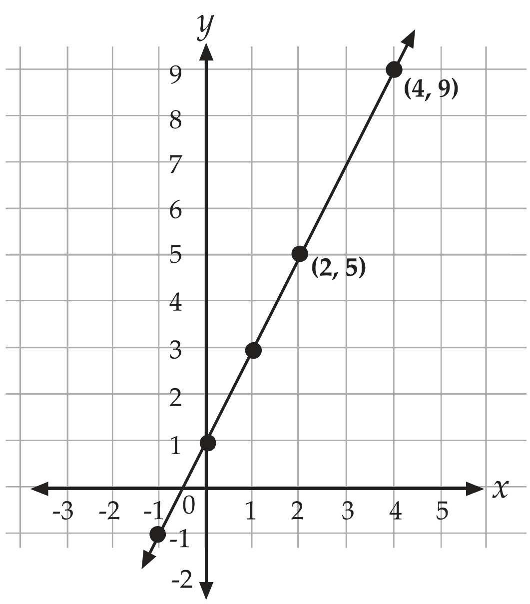
Another Way to Find the Slope of a Line

Here is another way to find the slope of a line. Pick any two points on a line. We will refer to the points as (x_1, y_1) and (x_2, y_2). The little numbers are called *subscripts* and they are just a way of letting the reader know that we are talking about *point 1* and *point 2*, and that they are *different* points. The subscripts mean that these are two distinct points of the form (x, y).
Note: The subscripts are *not* exponents.

Here is a *second* definition of slope:

$$\text{slope} = \frac{y_2 - y_1}{x_2 - x_1}$$

Example 2: Find the slope of this line.



Solution: Let (x_1, y_1) be (2, 5) and let (x_2, y_2) be (4, 9)

Substitute into the formula:

$$\text{slope} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{9 - 5}{4 - 2} = \frac{4}{2} = \frac{2}{1} \textit{ or } 2$$

Example 3: Find the slope of the line passing through (-7, 5) and (8, -2).

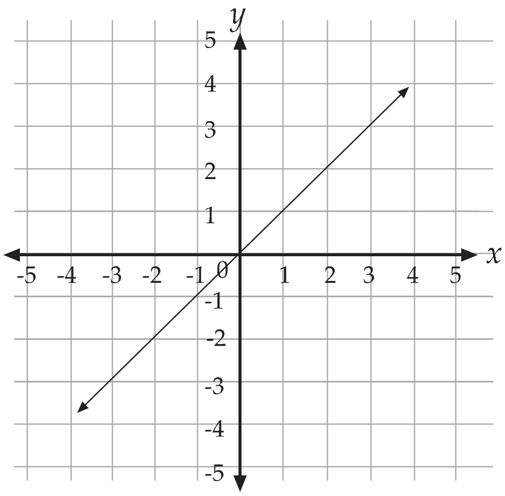
Solution: Let (x_1, y_1) be (-7, 5) and let (x_2, y_2) be (8, -2)

$$\text{slope} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-2 - 5}{8 - -7} = \frac{-7}{15}$$

Look to Be Sure—Positive and Negative Slopes

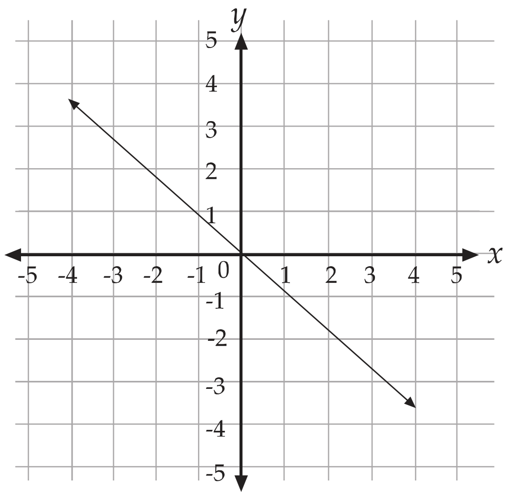
Look at the two graphs below. You can tell if the slope is positive or negative by looking. A line with a *positive slope* rises as the value of x increases. (See line 1 below.) A line with a *negative slope* falls as the value of x increases. (See line 2 below.)

line 1



positive slope

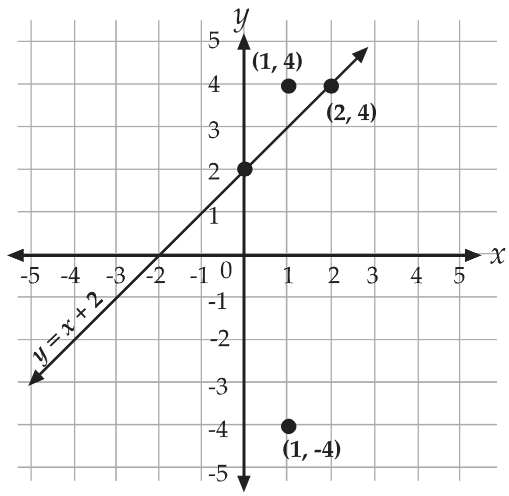
line 2



negative slope

Graphing Inequalities

Consider the line $y = x + 2$. We can think of the line separating the coordinate plane into three distinct parts.



- We have points above the line.
- We have points on the line.
- We have points below the line.

This awareness of the relative position of points with regard to the line is important when we need to solve **inequalities** like $y > x + 2$.


Remember: Inequalities state that one expression

—is greater than (>),
—is greater than or equal to (≥),
—is less than (<),
—is less than or equal to (≤), or
—is not equal to (≠)

another expression.

Let's investigate which points work in the above inequality. The *solution of an inequality with two variables* is an ordered pair of numbers that makes the inequality true. Are the points on the line solutions? For example, (2, 4) lies on the line. Does it make a true sentence when we substitute it into the inequality?

$y > x + 2$
$4 > 2 + 2$
$4 > 4$ This is false, so (2, 4) is *not* a solution!

If we investigate further, it turns out that *no* point *on* the line would work.

What about (1, -4)? This point was picked randomly, and it lies in the region below the line. Let's substitute into the inequality and see if this point is a solution.

$y > x + 2$
$-4 > 1 + 2$
$-4 > 3$ This also turns out to be *false*!

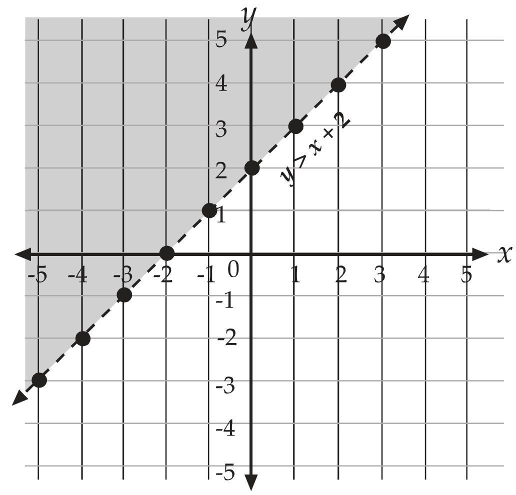
If we investigate further, it turns out that *no* point in the region *under* the line works.

Finally, we need to check and see if a point from the region above the line works. The point (1, 4) is such a point.

$y > x + 2$
$4 > 1 + 2$
$4 > 3$ This turns out to be true!

It also turns out that any point *above* the line works in the inequality.

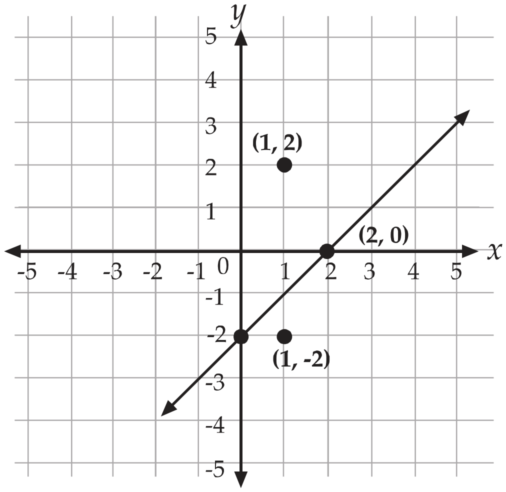
To show that these points are solutions, we shade the region above the line. Since no points on the line worked, we make the line *dotted*.



Example: Graph $y \leq x - 2$

Solution:

1. Graph the line $y = x - 2$. This will divide the coordinate plane into three regions—points *on*, *above*, and *below* the line.

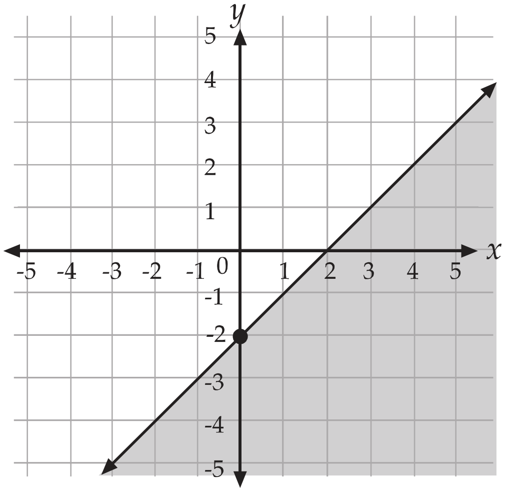


2. Pick any point above the line, on the line, and below the line.

3. Substitute the points into the inequality, and see which ones work:

above the line	**on the line**	**below the line**
(1, 2)	(2, 0)	(1, -2)
$y \leq x - 2$	$y \leq x - 2$	$y \leq x - 2$
$2 \leq 1 - 2$	$0 \leq 2 - 2$	$-2 \leq 1 - 2$
$2 \leq -1$	$0 \leq 0$	$-2 \leq -1$
This is false!	This is true!	This is true!

4. Make the line *solid*, since the test point (2, 0) was a solution. That means that *all* points on the line work. Shade below the line, since the test point (1, -2) was also a solution. That means that *all* points below the line are solutions.



When solving inequalities, you will either shade above the line *or* below the line. Never both.

Lesson Two Purpose

- Understand and explain the effects of addition, subtraction, multiplication, and division on real numbers. (MA.A.3.4.1)
- Apply special number relationships such as sequences to real-world problems. (MA.A.5.4.1)

Sequences

In real life, we frequently see numbers that form a **pattern (relationship)**. That is, numbers which appear in a specific order (or pattern) determine a **sequence**.

Street signs in cities and towns use numbers such as 1^{st} Avenue, 2^{nd} Avenue, 3^{rd} Avenue, etc. Likewise, house numbers form patterns with even numbers on one side of a street and odd numbers on the opposite side. If your house number is 1800, then your neighbor will most likely live at 1802 and the next neighbor at 1804, and so on.

When we shop for clothes and shoes, signs giving sizes are placed *in order* to help customers quickly locate needed sizes.

Example 1:

An apartment rents for $600 a month. Each year the monthly rent is expected to increase $25.00. What will the monthly rent be at the end of 4 years?

end of year	1	2	3	4
rent	$600	$625	$650	$675
	+25	+25	+25	

Notice that we could easily predict the monthly rent at the end of 5 years and at the end of 6 years.

Example 2:

Consider the numbers: 5, 6, 8, 11, 15, ?, ?, ? Can we predict the next 3 numbers? To see the pattern, we look at the difference between the numbers.

5 6 8 11 15 20 26 33

1 2 3 4 5 6 7

Example 3:

10^1	10^2	10^3	10^4	10^5
10	100	?	?	?

Using the pattern, what would be the value of 10^0? What would be the value of 10^6?

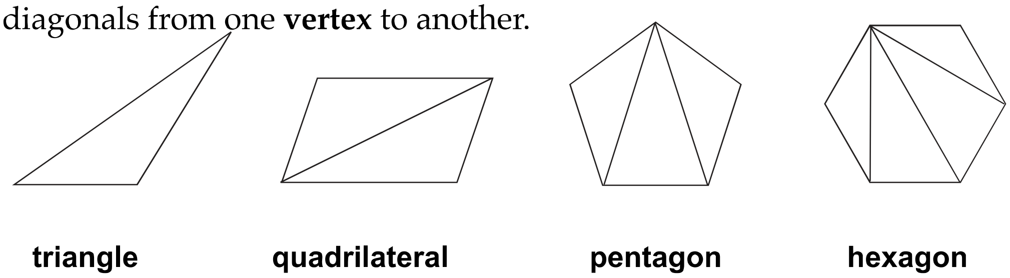
Answers:

$10^0 = 1$
$10^6 = 1{,}000{,}000$

Investigating Patterns in Polygons

Now that we have studied geometric shapes and line graphs, we are going to investigate some interesting topics that combine shapes, graphs, and patterns.

Recall that the *sum* of the **measures (m) of an angle (∠)** of one **triangle** is 180 **degrees (°)**. Consider the **polygons** below and notice that we have divided the polygons (with 4 or more *sides*) into triangles by drawing diagonals from one **vertex** to another.



The **quadrilateral** can be divided into 2 triangles, which means that the sum of its **angles** is 180 x 2, or 360°. How many triangles are formed in the **pentagon**? How many triangles are formed in the **hexagon**? Let's use a chart to summarize our information.

Number of Sides, Triangles, and Degrees in Polygons

Polygon	Number of Sides	Number of Triangles	Sum of Angle Measures
Triangle	3	1	1 x 180 = 180°
Quadrilateral	4	2	2 x 180 = 360°
Pentagon	5	3	3 x 180 = 540°
Hexagon	6	4	4 x 180 = 720°

Do you see a pattern in the angle measures? Each sum is a result of multiplying 180 times the *number of sides minus 2*.

Algebraically, if n represents the number of sides of the polygon, we discover the formula: $180(n - 2)$ or $(n - 2)180$. The formula can be used to calculate the *sum of the measures of the angles* of any polygon.

Also, if the polygon's angles are all **congruent (≅)**, or the same shape and size, we can find the measure of each angle of the polygon. Recall from Unit 4 that if all angles in a polygon have the same measure, we describe the polygon as an **equiangular polygon**—a polygon that has equal angles.

If a quadrilateral is *equiangular*, we can find the *measure of each angle* by dividing the sum of the angles by the actual number of angles. A quadrilateral has 4 sides and 4 angles. To calculate the measure of each angle, we use the formula:

$$\frac{180(n-2)}{n}$$

Since a quadrilateral has 4 sides and 4 angles, then $n = 4$.

$$\text{each angle in a quadrilateral} = \frac{180(4-2)}{4}$$

$$= \frac{180(2)}{4}$$

$$= \frac{360}{4}$$

$$= 90°$$

Lesson Three Purpose

- Determine probabilities using counting procedures, tables, and tree diagrams. (MA.E.2.4.1)
- Determine the probability for simple and compound events as well as independent and dependent events. (MA.E.2.4.2)

Tree Diagrams

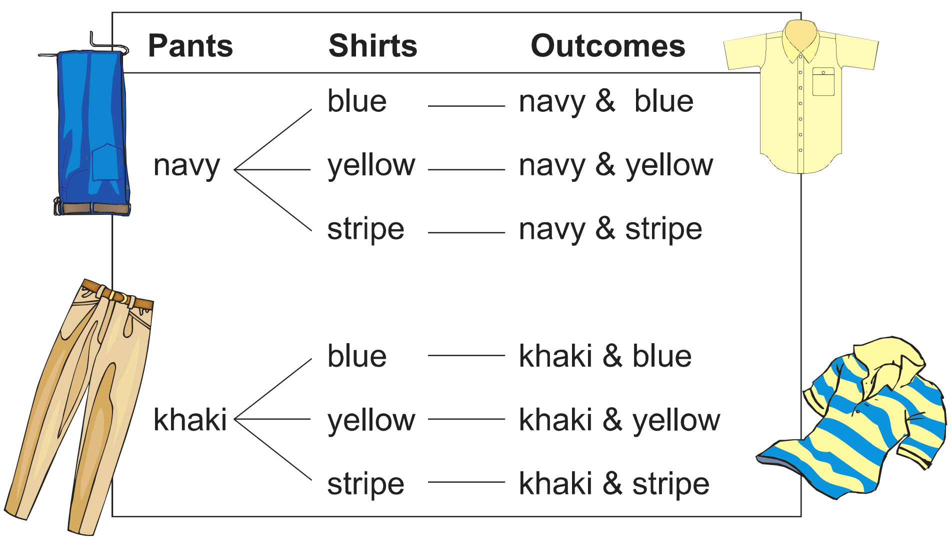
We can use a **tree diagram** to help us find the number of ways that one **event**, or possible *result* or **outcome**, can occur at the same time another *event* occurs. For example, we can consider different ways of combining garments to make outfits, or of looking at possible *outcomes* in games, or of ordering different combinations of items from a menu.

Example 1:

Suppose a school's colors are light blue and yellow. The school's uniforms allow khaki or navy pants to be worn with solid blue, solid yellow, or blue and yellow striped shirts. How many outfits are possible?

We can use a *tree diagram* to look at the various outfits. A tree diagram is a diagram that shows all possible outcomes of a given event.

School Uniform Color Outcomes

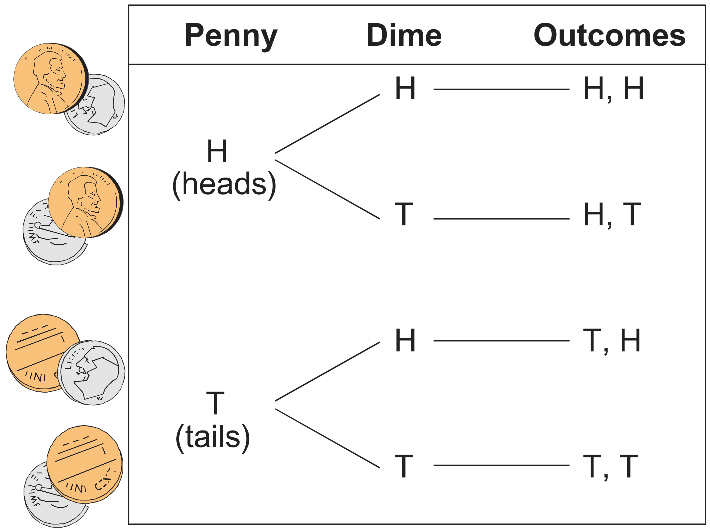


There are six possible outfits.

Example 2:

Suppose a penny and a dime are tossed at the same time. How many possible outcomes are there?

Coin Toss Outcomes



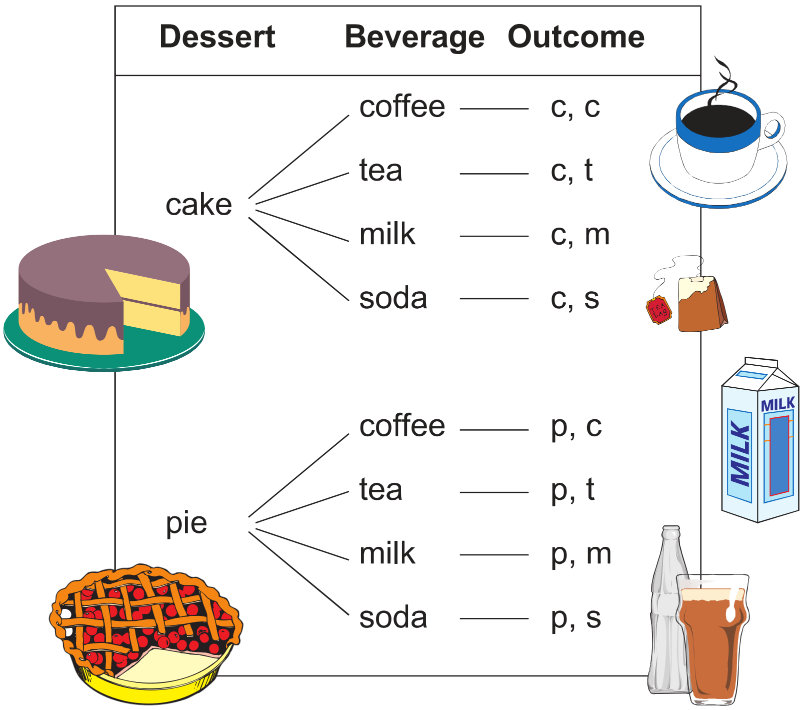
Notice that if *event A* has "a" outcomes and *event B* has "b" outcomes, then the total number of outcomes is (a)(b). The penny has 2 sides and the dime has 2 sides, so the outcomes are 2 x 2 = 4.

Look at the pants and shirts example with 2 choices for pants and 3 choices for shirts. The total number of outcomes was (2)(3) = 6.

Example 3:

Joe’s Snack Shop serves 2 desserts: cake and pie. They also serve 4 beverages: coffee, tea, milk, and soda. If you choose 1 dessert and 1 beverage, how many possible combinations are there?

Joe’s Snack Shop Outcomes



There are eight possible combinations.

Probability

At the beginning of a football game, a coin is tossed to determine which team will get the ball first. Let's say that Team A will get the ball first if the outcome is heads. Since the coin has two sides, it is **equally likely** that the coin will show heads or tails. *Equally likely* means that of the two possible outcomes—heads or tails—each has the same **probability** of occurring. *Probability* is the *ratio* of the number of favorable outcomes to the total number of outcomes.

$$\frac{\text{number of favorable outcomes}}{\text{total number of outcomes}}$$

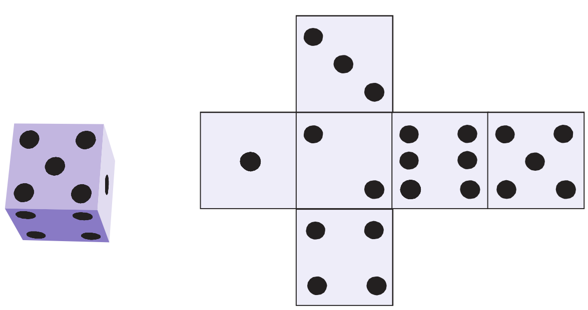
Let's look at the probability that Team A will get the ball first.

$$\text{probability of a a certain outcome} = \frac{\text{number of ways a certain outcome can occur}}{\text{number of possible outcomes}}$$

$$P(\text{heads}) = \frac{1}{2} \quad \frac{\text{one outcome of heads can occur}}{\text{there are 2 equally likely outcomes, heads or tails}}$$

Remember: *P* means probability; *P*(heads) refers to the probability of heads occurring.

Many games are played by rolling a die. Each die is in the shape of a **cube**. Recall from Unit 4 that a *cube* has six congruent sides and that each side is a square. Look at the die and the *flattened* **net** of the die below. The *net* is the plan which can be used to make a model of a solid. The net below forms a cube-shaped die.



If you rolled the die, there would be 6 equally likely outcomes. That means that each number has the same chance of landing face up.

Examples:

- Find the probability of rolling a 5.

$$P(5) = \frac{\text{number of ways of rolling a 5}}{\text{number of possible outcomes}}$$

$$P(5) = \frac{1}{6}$$

- Find the probability of rolling an odd number.

$$P(\text{odd}) = \frac{\text{number of odd numbers}}{\text{number of possible outcomes}}$$

$$P(\text{odd}) = \frac{3}{6} \quad \frac{\text{odd numbers are 1, 3, and 5}}{\text{there are 6 possibilities}}$$

$$P(\text{odd}) = \frac{1}{2} \quad \text{always simplify fractions}$$

- Find the probability of rolling either a 3 or a 4.

$$P(3 \text{ or } 4) = \frac{2}{6}$$

$$P(3 \text{ or } 4) = \frac{1}{3}$$

- Suppose you roll a die. What is the probability of rolling a 7?

$$\frac{\text{number of ways a 7 can occur:}}{\text{number of possible outcomes:}} = \frac{0}{6} = 0$$

$$P(7) = 0$$

- What is the probability of rolling a number from 1 through 6?

$$\frac{\text{number of ways a 1, 2, 3, 4, 5, or 6 can occur:}}{\text{number of possible outcomes:}} \quad \frac{6}{6} = 1$$

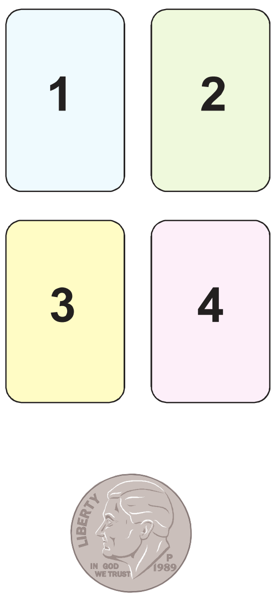
$$P(1 \text{ through } 6) = 1$$

> **Note:** An outcome that *will definitely happen* has a probability of 1. An outcome that *cannot happen* has a probability of 0.

Probability of Independent Events

Think about first selecting a card and then tossing a coin. Since the outcome from selecting a card in no way affects the outcome of tossing the coin, these activities are called **independent events**. With *independent events*, the first event does *not* affect the outcome of the second event.

The tree diagram below shows all possible outcomes.



Independent Card and Coin Outcomes

Cards	Dime	Outcomes
1	H	1, H
	T	1, T
2	H	2, H
	T	2, T
3	H	3, H
	T	3, T
4	H	4, H
	T	4, T

There are eight possible outcomes.

Examples:

- The probability of drawing a 3 is 1 out of 4 chances or

  $P(3) = \frac{1}{4}$.

- The probability of tossing heads (H) is 1 out of 2 chances or

  $P(\text{H}) = \frac{1}{2}$.

- The possibility of drawing a 3 and then tossing heads is 1 out of 8 or

  $P(3 \text{ and H})$ is $\frac{1}{8}$ because $P(3) = \frac{1}{4}$ and $P(\text{H}) = \frac{1}{2}$ so

  $P(3 \text{ and H}) = P(3)$ times $P(\text{H})$ or $\frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$.

Below is the formula for finding the probability of independent events.

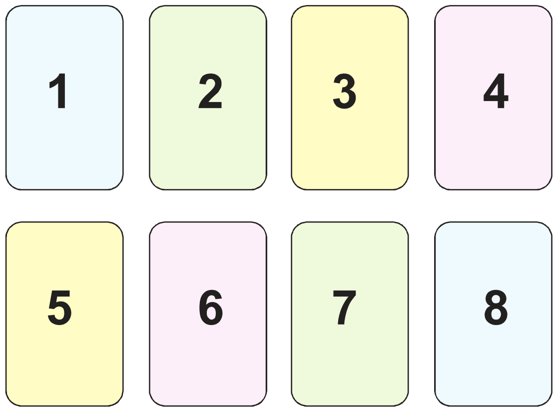
Probability of Independent Events

The probability of two *independent* events (*A* and *B*) both occurring can be found by multiplying the probability of the first event (*A*) by the probability of the second event (*B*).

$$P(A \text{ and } B) = P(A) \times P(B)$$

Probability of Dependent Events

Suppose these cards are shuffled and placed face down. What is the probability of drawing a 5?



We know that $P(5)$ is $\frac{1}{8}$ because the 5 is 1 card out of 8 cards. Suppose you draw the 5 and do *not* replace the card, and then draw a second card. What is the probability that the second card drawn will be a 2?

Since the first card was not replaced, the probability of drawing a 2 is $\frac{1}{7}$ because the 2 is 1 card out of the 7 cards left. The outcome of drawing the first card *affects* the outcome of drawing the second card from the remaining cards. Therefore, we describe the events of drawing a 5 and then drawing a 2 as being **dependent events**. With *dependent events*, the first event *affects* the outcome of the second event.

$$P(5, 2) = \frac{1}{8} \text{ x } \frac{1}{7}$$
$$= \frac{1}{56}$$

So the probability of drawing a 5 and then drawing a 2 is $\frac{1}{56}$, or one chance out of 56.

The Probability of Dependent Events

The probability of two *dependent* events (*A* and *B*) both occurring can be found by multiplying the probability of the first event (*A*) and the probability of the second event (*B*) after the first event (*A*) occurs.

$$P(A \text{ and } B) = P(A) \times P(B \text{ following } A)$$

Lesson Four Purpose

- Understand and use the real number system. (MA.A.2.4.2)
- Interpret data that has been collected, organized, and displayed in charts, tables, and plots. (MA.E.1.4.1)
- Calculate measures of central tendency (mean, median, and mode) and dispersion (range) for complex sets of data and determine the most meaningful measure to describe the data. (MA.E.1.4.2)
- Explain the limitations of using statistical techniques and data in making inferences and valid arguments. (MA.E.3.4.2)

Measures of Central Tendency—Mean, Median, and Mode

The **mean**, **median**, and **mode** are sometimes called **measures of central tendency**. *Measures of central tendency* describe how *data* is *centered*. Each of these measures describes a set of data in a slightly different way.

- The **mean (or average)** is the sum of the data divided by the number of items.

  When the data is centered around the *mean*, this data is considered an appropriate measure of central tendency. The mean can be distorted by an *extreme* value, a value that is much greater than or less than the other values.

- The **median** is the middle item when the data is listed in numerical order. If there is an even number of items, the median is the average of the two middle numbers.

  When there is an extreme value in a set of data that distorts the mean, the *median* is an appropriate measure of central tendency.

- The **mode** is the item that appears most often. There can be more than one mode. There is *no* mode if each item appears only once.

  When data cannot be averaged (to find a mean) or listed in numerical order (to find a median), the *mode* is the appropriate measure of central tendency.

The **range** (of a set of numbers) is the difference between the highest and lowest value in a set of data. The *range* can help you decide if differences among the data are important.

Example 1:

Consider Larry's math test scores for the last 9 weeks:

0, 55, 70, 80, 80, 82, 85, 88, 100

Notice that the grades are in order from lowest to highest. Since there are three measures of central tendency and a range to use, we will consider each one to arrive at a fair grade for Larry.

- The *mean* (or average) is the sum of the numbers in the **set** divided by how many numbers are in the *set*. The mean is what most people think of when they think of finding an *average*. The set is the collection of numbers.

$$\frac{0 + 55 + 70 + 80 + 80 + 82 + 85 + 88 + 100}{9} = \frac{640}{9} = 71 \text{ rounded to nearest whole number}$$

Larry's mean or average grade would be 71.

- The *mode* is the *number that occurs most often* in a set of data. Sometimes there is no mode, and sometimes there are several modes.

  0, 55, 70, **80**, **80**, 82, 85, 88, 100

Larry's mode would be 80 because it appears twice.

- The *median* is the *number in the middle*. To find the median, the numbers have to be placed in numerical order.

  0, 55, 70, 80, **80**, 82, 85, 88, 100

In this case 80 is the median because it is the number in the middle.

Besides these three measures of central tendency, we can also consider the range of Larry's grades.

The *range* can be reported in one of two ways.

- The range can be reported as the *lowest value to the highest*.

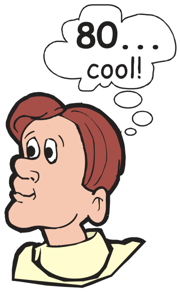
  The range of Larry's grades is 0 to 100 because 0 is the lowest value and 100 is the highest value.

- The range can also be reported as the **difference** between the *highest and lowest values*.

  The range of Larry's grades is 100 because 100 is the difference between the smallest number (0) subtracted from the largest number (100) in the set.

Let's review what happened when we used the three measures of central tendency and the range:

mean 71
mode 80
median 80
range 100



Probably the fairest grade to give Larry for the 9 week grading period would be an 80. Most of the time Larry appears to be a "B" student, yet his average (mean) was a 71. The bottom two grades of 0 (for the test he missed) and 55 distorted the average. This is something that you need to be aware of when you read about averages in the paper.

Example 2:

Let's look at a second example. Below are the weights of 8 tenth grade girls:

90, 100, 110, 120, 124, 130, 132, 150

We will find the three measures of central tendency.

- *mean* (average)

  $$\frac{90 + 100 + 110 + 120 + 124 + 130 + 132 + 150}{8} = \frac{956}{8} = 120 \text{ rounded to nearest whole number}$$

- *mode* (the number which appears the most)

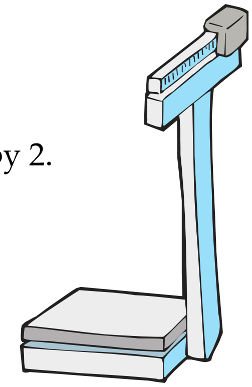
  Notice that all the numbers are different. Therefore there is no mode.

- *median* (the middle)

  90, 100, 110, 120, ↓ 124, 130, 132, 150

  To find the median, add 120 and 124, then divide by 2.

  $$\frac{120 + 124}{2} = \frac{244}{2} = 122$$



Summary: mean 120
mode .... not one
median 122

Both the mean and the median appear to be good representations of the data.