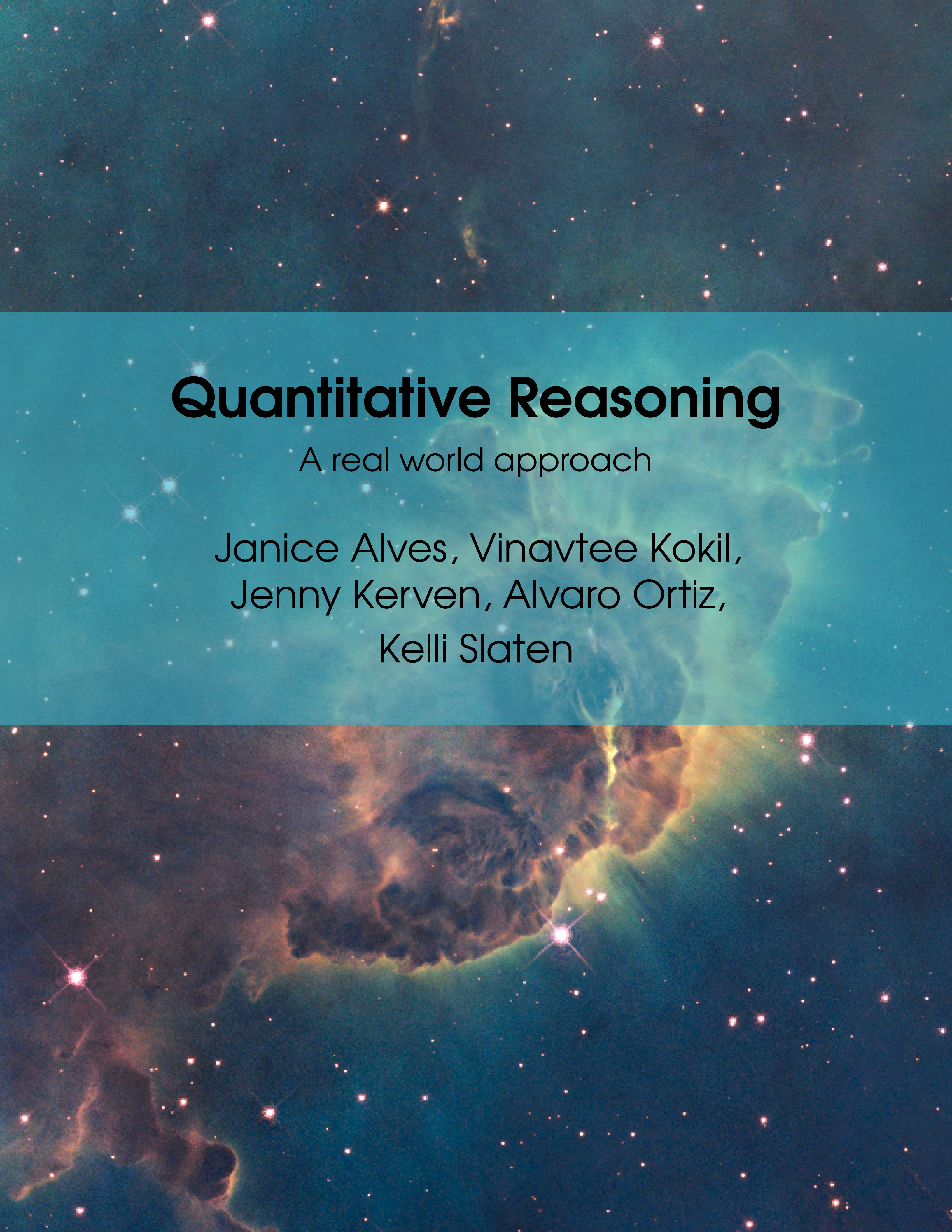


# Quantitative Reasoning

A real world approach

Janice Alves, Vinavtee Kokil,
Jenny Kerven, Alvaro Ortiz,
Kelli Slaten

Typeset in $\LaTeX$.

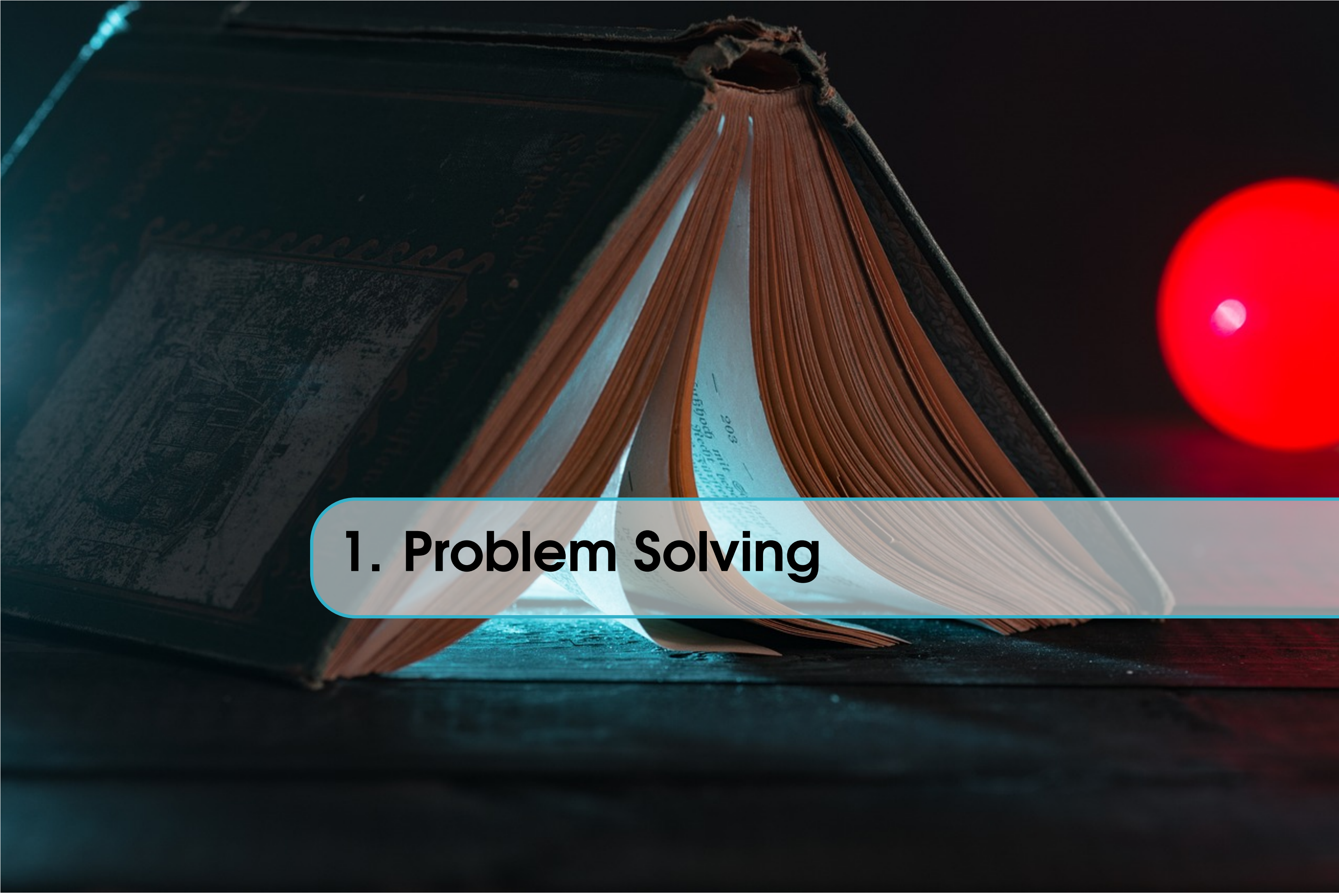


This book was created under the funding of the Affordable Learning Grant 2021-2022.

The date is the main identifier of version. The edition number is raised only if there have been substantial changes. Edition number started at 1, that is, version 1.0, as it was not kept track of before.

*First release, August 2021*


# Contents



# 1. Problem Solving

This first chapter will review some of the mathematical ideas that we use daily, even though we may not realize that we are performing mathematical computations. Everyone is familiar with decimals and fractions, but do you know every day that you go to the store and purchase an item, you are using a fraction or decimal converted into a percent? When people want to buy things, the computer or a phone is a valuable tool. This saves time and gas from traveling around town looking for these items, but it helps you compare and find the best deal. And who doesn't want to save some money? This is a mathematical concept called a rate.

The concept of rates can be extended into proportions; for example, you would like to take a trip somewhere. You fill up the gas tank, and off you go, but you will need to use your math skills to determine how many miles you can drive before you need to fill your tank up again. It might be beneficial to calculate how long it's going to take you to get to your final destination just in case someone is waiting for a home-cooked meal for you.

The last thing in this chapter that will be reviewed is geometry. It's not just about shapes and formulas but how they are used. Look around your house and see all kinds of geometrical forms. There are rectangles, squares, semi-circles, and area and perimeter.

Someday, if you haven't already, you will want to paint a room, and you will need to determine how much paint you will need.

You may want to carpet your floor, so you will need to know the amount of carpet. Maybe you will want to put a garden in the backyard, and you will need to see the amount of soil you will need to build up the bed. Don't forget the fence around it so that the animals won't get in to eat the vegetables. As you can see, math is used all the time. It's essential to know how to use math correctly and give it a chance because it's fun!

The contents of this chapter are derived from works such as Math in Society (see [Lip17]), The Open Stax Pre-Algebra (see [Mar20]), as well the Pre Algebra Textbook from College of the Redwoods (see [Wag09]) and Math for Liberal Arts Students from Darlene Diaz (see [Dia17]).

## 1.1 Decimals

We can see decimals everywhere we look. From shopping at the grocery store to buying gas. From the road markers on the road indicating specific locations to the times in racing competitions.


Figure 1.1: Decimals numbers arise in many situations of our daily life, like when we are paying for gasoline (photo credit: pixelnaiad).

A decimal answer gives us more of a precise solution than a whole number. It is an easier way of writing a mixed number fraction. The left side of the decimal indicates the integer part. The numbers to the right of the decimal point signal the fractional part of the decimal. Which are in powers of 10. When you read the fractional part, you will have a "th" at the end of the word, such as tenth, hundredth, thousandth, etc.

The easiest way to read a decimal is to read the decimal part as a fraction. Suppose we had 0.4 grams of yogurt in a cup. We would say, "4 tenths of a gram of yogurt," as the 4 is in the tenths place.

The table below shows the relationship between the whole numbers and the decimals. It also shows how the decimals are related to fractions.

| Whole Number | | Name |
|---|---|---|
| 1 | | One |
| 10 | | Ten |
| 100 | | One hundred |
| 1,000 | | One thousand |
| 10,000 | | Ten thousand |

| Decimal | Fraction | Name |
|---|---|---|
| 0.1 | $\frac{1}{10}$ | One tenth |
| 0.01 | $\frac{1}{100}$ | One hundredth |
| 0.001 | $\frac{1}{1,000}$ | One thousandth |
| 0.0001 | $\frac{1}{10,000}$ | One ten-thousandth |

### 1.1.1 Reading and Writing a Number in Decimal Form

**■ Example 1.1 — Decimal to words.** .

These examples show how a number that is written as a decimal form should be read and written.

(a) 9.4 is stated as, "Nine and four tenths".

(b) 87.49 is stated as, "Eighty-seven and forty-nine hundredths".

(c) 594.236 is stated as, "Five hundred ninety-four and two hundred thirty-six thousandths".

■

You should notice that the word "and" was read as the decimal place every time we wrote the decimal. "And" should never be written or said except when it represents the decimal point.

When you look at the table with the decimals and fractions, you should observe that the number of zeros in the fraction's denominator is always the same as the number of digits past the decimal place in the original number given. For example, the number one-thousandth has three numbers after the decimal place. Therefore, if you write this as a fraction, there would be three zeros in the fraction's denominator. Not all of your answers in fraction form will result in denominators of 10, 100, 1,000, etc. because you will permanently reduce your fraction into the simplest form.

**Quick Practice - Decimals: Read and Write**

**Exercise 1.1.1 — Decimal to words.**

Write the name of the number or write the number, as indicated in the directions.

(a) Write the name of the number 2.45.

(b) Write the name of the number 12.791.

(c) Write the name of the number 0.094.

(d) Write the name of the number 0.000208.

(e) Write the number number six and seventeen hundredths.

(f) Write the number two hundred twenty and seven hundred and sixty thousandths.

(g) Write the number twenty-four thousandths.

(h) Write the number nine hundred thirty-six ten thousandths.

**Solutions:**

(a) Two and forty-five hundredths

(b) Twelve and seven hundred ninety-one thousandths

(c) Ninety-four thousandths

(d) Two hundred eight millionths

(e) 6.17

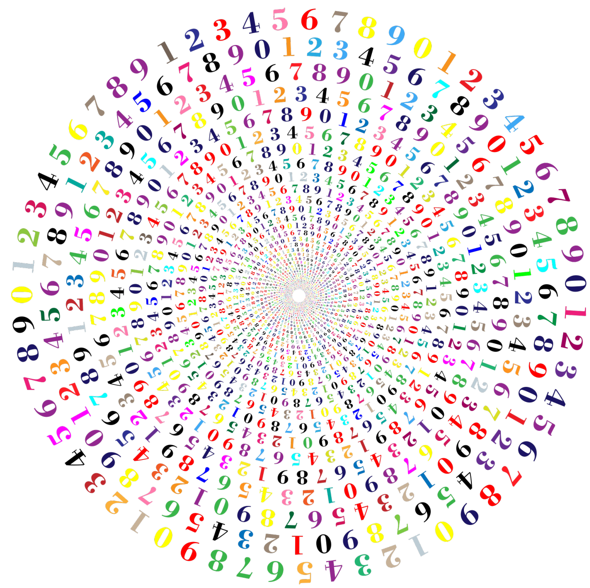
(f) 220.760

(g) 0.024

(h) 0.0936

### 1.1.2 Comparing Decimals and Rounding Decimals

There are times where we would like to compare decimals. We can do this by comparing digits in each place as we move from left to right, place by place.



**■ Example 1.2 — Part and fraction.** .

Suppose we wish to compare the decimal numbers 5.234 and 5.2357. Align the decimals vertically. Add zeros to the end of the number with the fewest decimal places so that the numbers have the same number of decimal places. These zeros do not change the value of the number. Begin at the left side of the numbers in the same place value. Compare each of these numbers. Once you find two different digits in the same place, the decimal number with the greatest digit in this place value is the more significant number.

The figure below shows how to compare the numbers, moving from left to right.

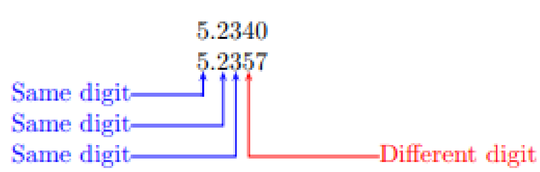



Figure 1.2: 5.2357 > 5.234

When rounding decimals, there is a rounding digit and a test digit. ■

**Definition 1.1.1** The digit in the place to which we wish to round is called the **rounding digit** and the digit that follows on its immediate right is called the **test digit**.


**Example 1.3** If we want to round the decimal number 12.254 to the nearest hundredth, then the rounding digit is five, and the test digit is 4.

Then follow these rules for rounding:

1. If the test digit is greater than or equal to 5, add 1 to the rounding digit and remove all of the digits to the right of the rounding digit.
2. If the test digit is less than 5, remove all of the digits to the right of the rounding digit.

The figure below shows the process of rounding 12.254 to the nearest hundredth. ∎

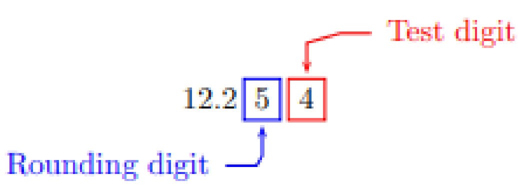



Figure 1.3: 12.254 becomes 12.25 when rounded to the nearest hundredth

In the above example, what would the number be if 12.254 was rounded to the nearest tenth place?

The rounding digit would be the two, and the test digit would be the 5. Since the test digit is five or greater, we would add 1 to the rounding digit and remove the rest of the numbers to the right of the rounding digit. Therefore, the number would be 12.3.

### Quick Practice - Decimals: Comparing and Rounding

**Exercise 1.1.2** For a - c, determine which statement is correct. For d - f, round the number to the specified place value.

(a) $0.30387617 < 0.3036562$ or $0.30387617 > 0.3036562$

(b) $8.5934 < 8.554$ or $8.5934 > 8.554$

(c) $18.62192 < 18.6293549$ or $18.62192 > 18.6293549$

(d) Round 79.369 to the nearest hundredth.

(e) Round 89.3033 to the nearest thousandth.

(f) Round 53.967 to the nearest tenth.

**Solutions:**

(a) $0.30387617 > 0.3036562$

(b) $8.5934 > 8.554$

(c) $18.62192 < 18.6293549$

(d) 79.37

(e) 89.303

(f) 54.0

### 1.1.3 Accuracy and Precision

Accuracy means how close your measured value is to the exact value of the item you are working with. Precision means how close the measured values are to each other.

■ **Example 1.4** I have asked two groups of 3 students in my class to guess what number I am thinking of, between 1 and 25. Determine if the answers are precise, accurate, both, or neither.

The number I was thinking of was 17.

Group one's answers were 15, 19, and 21, and group, two's answers, were 31, 33, 30.

The first group's answers were close to my exact value of 17. The second group's numbers were not close to the actual value but their measured values. Therefore, group one's answers are more accurate, and group two's answers are more precise. ■

### 1.1.4 Approximation

Approximating a number is similar to rounding, except we want to round the number quickly to make it easier to perform calculations. We are not looking for an exact answer but a reasonably accurate one. Another objective of approximating is that we want to be able to calculate without using a calculator.

■ **Example 1.5** You are on a road trip of 256.27 miles, and it takes 4.3 hours to drive. How many miles per hour do you drive?


If we rounded to the nearest tens place, the miles would be 256, and the hours would be 4. This calculation would not be straightforward without a calculator so let's approximate instead. Let's use 300 miles and 5 hours. The calculation would be

$$\frac{300}{5} \approx 60 \text{ miles per hour.}$$

If you used your calculator, you would find that the answer is 59.6 miles per hour, which is the same answer without approximating. Not all answers will be the same, but they should be close to each other. ■

When we use an approximation, a round-off error is likely to occur. This happens because we are not using exact values. After all, we have rounded all our numbers to make our calculations easier. As these rounded numbers are used throughout our estimates, and the answers are rounded, the final answer can be quite significantly different from the answer that would have been calculated if the exact values had been used.

Some of these errors can have devastating effects. For example, The Patriot missile defense system used during the Gulf War was ineffective due to round-off error (Skeel 1992, U.S. GAO 1992). The error was approximately 0.3433 seconds, and the Patriot missile defense system could not target the Iraqi Scud missile. Unfortunately, it detonated in a barracks, and 28 people were killed. `https://mathworld.wolfram.com/RoundoffError.html`.

### 1.1.5 Percent Error

We can find the percent error to determine how accurate a measurement is, being small or large; this error will indicate how accurate the measured value is to the actual value.

If the error is small, less than 5%, your measurement is very accurate to the actual value. If the error is significant, over 5%, your measurement is not precise to the real value. To avoid any confusion, we will use the absolute value of the difference between the measured and actual value so that the percent will be positive.

**Definition 1.1.2** To find the **Percent Error**, you must be given two quantities, the measured value and actual value. Then use the formula to calculate the error:

$$\frac{\|\text{ measured value - actual value }\|}{\text{actual value}} * 100$$

■ **Example 1.6** Find the percent error, rounded to the tenth place, and determine the accuracy.

(a) For a wedding reception, the estimated amount of people to attend was 375. The actual amount of people that attended was 325. The percent error is:

$$\frac{\|\ 375 - 325\ \|}{325} * 100 = 0.154 * 100 = 15.4\%$$

The estimate was not very accurate since it was not close to the actual value. Therefore, the bride and groom may have lots of food to take home.

(b) Oreo, my domestic cat went to the veterinarian for a check-up. I estimated his weight at 10 pounds. The vet weighed him and the weight was 10.4 pounds. The percent error is:

$$\frac{\|\ 10 - 10.4\ \|}{10.4} * 100 = 0.038 * 100 = 3.8\%$$

The estimate was low so it's considered to be accurate. The estimated value is close to the actual value.

■

### 1.1.6 Significant Digits

To determine how precise a number is, we can use significant digits. The more significant digits present in the number, the more accurate the measurement will be.

To count the number of significant digits in a number, follow these rules:

1. Look for a decimal. If there is a decimal, start on the left side of the number. If a decimal is not present, start on the right side of the number.

2. Begin to count the numbers that are not zero.

3. Continue counting until you get to the end (or the front) of the number.

■ **Example 1.7** Determine the number of significant digits.

(a) 670 - There is no decimal, so start on the right side. As we move left, the first number is a 0, so we cannot start counting. Move to the next number, and we have a seven, so you can begin counting and do not stop until you reach the front of the number. Therefore, 670 has two significant digits, 7 and 6.

(b) 670. - There is a decimal, so start on the left side. As we move right, the first number is a seven, so we can start counting until we reach the end of the number. Therefore, 670. has three significant digits, and they are 6, 7, and 0.

The difference between a and b is the decimal point. We started on the right side of the number but could not count the zero. In b, we started on the left side of the number and could count every digit.

(c) 9.005 - There is a decimal, so start on the left side. As we move right, the first number is a nine, so we can start counting until we reach the end of the number. Therefore, 9.005 has four significant digits, and they are 9, 0, 0, and 4.

(d) 0.005084 - There is a decimal, so start on the left side. As we move right, the first three numbers are 0, so we cannot start counting until we get to 5. Then we can continue counting until we reach the end of the number. Therefore, 0.005084 has four significant digits, and they are 5, 0, 8, and 4.

■

Instead of rounding, we may be asked to calculate some numbers and write the final answer to the appropriate number of significant digits. This will be different if we are adding and subtracting or multiplying and dividing.

For adding and subtracting, we will round to the fewest places to the right of the decimal.

For multiplying and dividing, we want to count the number of significant digits in each of the numbers used in the calculation. We will write the answer with the least amount of significant digits. We will use the rounding rules when determining the final solution.


■ **Example 1.8** Calculate and write the answer to the appropriate number of significant digits.

(a) Multiply 13 and 45.6: 13 has 2 significant digits and 45.6 has 3 significant digits, so your final answer will have 2 significant digits. Therefore,

$$13 * 45.6 = 592.8 \approx 590$$

(b) Divide 49.836 by 0.0103: 49.836 has 5 significant digits and 0.0103 has 3 significant digits, so your final answer will have 3 significant digits. Therefore,

$$\frac{49.836}{0.0103} = 4838.44660... \approx 4840$$

(c) Add 3.96 to 4.728: 3.96 has 2 digits past the decimal place and 4.728 has 3 digits past the decimal place. Therefore, the final answer will have 2 digits past the decimal place.

$$3.96 + 4.728 = 8.688 \approx 8.69$$

■

### Quick Practice - Percent Error and Significant Digits

**Exercise 1.1.3** For a - c, find the percent error and determine if the error indicates whether the measured value is accurate. Round each answer to the nearest tenth place or thousandths. For d - f, perform the calculations and write the answer with the correct number of significant digits.

(a) My car needed repairs and I estimated the cost to be $95. The bill was for $135.

(b) The bank took my rolled coins and told me that one roll only had 48 pennies instead of the required 50 pennies.

(c) Students in science class had to measure the length of a plant. The actual length was 31.4961 inches. Sam used a yard stick and measured the plant to be $31\frac{3}{4}$ inches.

(d) 19.8743 - 6

(e) 12.439 + 8.365 + 0.719 + 13.8

(f) $\frac{16.1 * 45.2207}{0.6}$

**Solutions:**

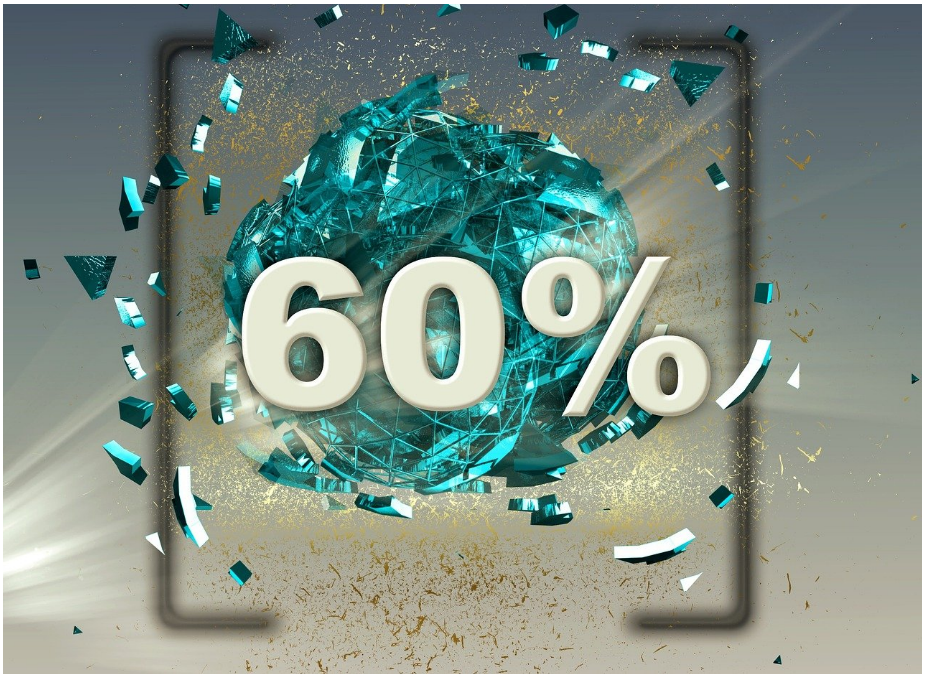
(a) 29.6%, the estimate was not very accurate.

(b) 4.2%, the estimate was fairly accurate.

(c) 0.8%, the estimate was very accurate.

(d) 14 (e) 35.3 (f) 1,000

## 1.2 Percents




Let's start by showing some ways that percents are used in every day life:

1. Shopping: paying tax and buying something on sale or with a coupon.
2. Banking: depositing money, taking out a loan, owning a credit card.
3. Nutrition: checking labels for nutrition and calorie counts.
4. Electronics: recharging your computer or phone battery before it no longer functions.
5. School: Grade distribution.
6. Paychecks: determining your taxes

**Definition 1.2.1** A **percent** is a ratio whose denominator is 100.

$$r\% = \frac{r}{100}$$

*percent* means "per 100," or "how many out of 100." You use the symbol % after a number to indicate percent.

Percentages will appear in multiple ways given a situation. Next, we see some examples of the different representations.

### 1.2.1 Percents to Fractions

To change a percent to a fraction:

1. Write the percent as a ratio with the denominator 100.
2. Simplify the fraction if possible.

■ **Example 1.9** Change 40% and 125% to a fraction:

$$125\% = \frac{125}{100} = \frac{5}{4}$$

$$40\% = \frac{40}{100} = \frac{2}{5}$$

**Calculator Keystroke:**

$40 \div 100$ enter (then change to a fraction using the instructions at the end of this section). ■

If the percent has a decimal, we can still put that number over 100. When the percent has a fractional part, we will change that percent into an improper fraction.

■ **Example 1.10** Change 24.5% and $33\frac{1}{3}\%$ to a fraction:

$$24.5\% = \frac{24.5}{100} = \frac{49}{200}$$

To change $33\frac{1}{3}\%$ to an improper fraction, take the whole number (33), multiply it by the denominator (3) and add the numerator (1). So you have $33 * 3 + 1 = 100$. The number goes over the denominator of 3 so your improper fraction becomes $\frac{100}{3}$.

$$33\frac{1}{3}\% = \frac{\frac{100}{3}}{\frac{100}{1}} = \frac{100}{3} \div \frac{100}{1} = \frac{100}{3} * \frac{1}{100} = \frac{1}{3}$$

**Calculator Keystroke:**

$(100 \div 3)$ $\div 100$ enter (then change to a fraction using the instructions at the end of this section.) ■

### 1.2.2 Fractions to Percents

To change a fraction to a percent:

1. Convert the fraction to a decimal. If the fraction is a mixed number, it will need to be changed to an improper fraction. To convert the fraction to a decimal, you will divide the numerator by the denominator.

2. Convert the decimal to a percent; this means you will multiply the decimal by 100. Then add the % sign.

**Example 1.11** Change

$$\frac{11}{8},\ 2\frac{1}{5},\ \text{and}\ \frac{2}{9}$$

to a percent:

(a) First we divide to obtain a decimal.

$$\begin{aligned}\frac{11}{8} &= 1.375\\ &= 1.375 * 100 \ \ \text{(Convert the decimal to percent)}\\ &= 137.5\%\end{aligned}$$

(b) First we make this into an improper fraction.

$$\begin{aligned}2\frac{1}{5} &= \frac{11}{5}\\ &= 2.2 \ \ \text{(Divide to Obtain a decimal)}\\ &= 2.2 * 100 \ \ \text{(Convert the decimal to percent)}\\ &= 220\%\end{aligned}$$

(c) First we divide to obtain a decimal.

$$\begin{aligned}\frac{2}{9} &= 0.\overline{2}\\ &= 0.\overline{2} * 100 \ \ \text{(Convert the decimal to percent)}\\ &= 22.2\%\end{aligned}$$

Since we multiply the decimal by 100, you should notice that the decimal is being moved to the right by two places. ■

### 1.2.3 Percents to Decimals

To change a percent to a decimal:

1. Write the percent as a ratio with the denominator 100 and remove the % sign.

2. Convert the fraction to a decimal by dividing the numerator by the denominator.

■ **Example 1.12** Change 6%, 135%, and 12.5% to a decimal:

$$6\% = \frac{6}{100} = 0.06$$

$$135\% = \frac{135}{100} = 1.35$$

$$12.5\% = \frac{12.5}{100} = 0.125$$

Since we are dividing the percent by 100, you should notice that the decimal is moved to the left by two places. ■

■ **Example 1.13** Transform the following numbers into a fraction and then to decimal.

(a) 0.0025%

(b) $-1.25\%$

**Solution:**

(a) We have a decimal percentage in this example. This may seem confusing, but it makes the number easier to understand once we convert it from the percentage. As you can see, this is a tiny number.

$$\begin{aligned} 0.0025\% &= \frac{0.0025}{100} \quad \text{(Divide over hundred to convert to fraction)} \\ &= 0.000025 \quad \text{(Convert the percent to decimal)} \end{aligned}$$

(b) In this example, we have a negative percentage. However, this will not change how we do the conversion. The final answer will be negative since we divided a negative number by a positive number.

$$\begin{aligned} -1.25\% &= \frac{-1.25}{100} \quad \text{(Divide over hundred to convert to fraction)} \\ &= -0.0125 \quad \text{(Convert the percent to decimal)} \end{aligned}$$

■

### 1.2.4 Summary on calculating percents, fractions, and decimals

Percents, fractions, and decimals are all common ways to represent fractional amounts.

- To change any percent into a fraction, write the percent as a ratio with the denominator of 100 and remove the % sign from the number in the numerator. If the percent is written as a mixed number, convert it to an improper fraction before it is put into the numerator. This will force you to have a "complex fraction."

- To change a fraction into a decimal, divide the numerator by the denominator. If the fraction is a mixed number, convert it to an improper fraction before dividing.

- To change a decimal into a percent, multiply the number by 100 and add the % sign or move the decimal two places to the right. Additional zeros may need to be added to the number. For example, $0.6 = 60\%$. The zero after the six had to be added.

- To change a percent into a decimal, divide the number by 100 and remove the % sign or move the decimal two places to the left. Again, additional zeros may need to be added to the number. For example, $3\% = 0.03$. The zero before the three had to be added.
  When a number does not show a decimal, it is understood to be at the end of the number.

**Quick Practice - Percents:**


**Exercise 1.2.1** Make the indicated conversions and leave answers in simplified form.

(a) 0.17% to a decimal.

(b) 8.6 to a percent.

(c) 11.76% to a fraction.

(d) $3\frac{6}{7}$ to a decimal.

(e) 5.9 to a fraction.

(f) $8\frac{5}{8}$ to a percent.

(g) When changing a decimal into a percent, you can (multiply or divide) by 100 or move your decimal to the (right or left)?

(h) When changing a percent into a decimal, you can (multiply or divide) by 100 or move your decimal to the (right or left)?

**Solutions:**

(a) 0.0017

(b) 860%

(c) $\frac{147}{1250}$

(d) 3.86

(e) $\frac{59}{1000}$

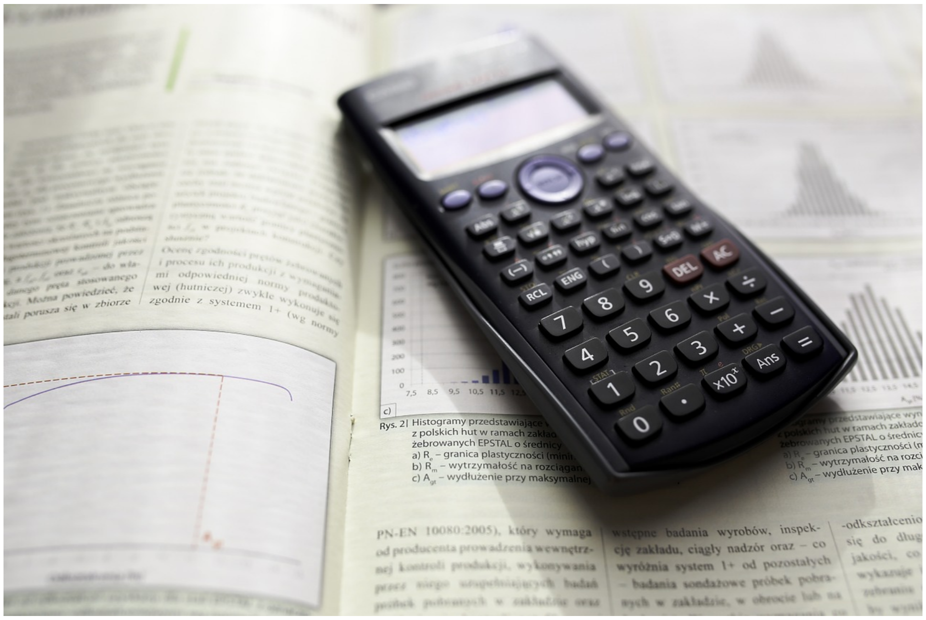
(f) 862.5%

(g) Multiply, right

(h) Divide, left

### 1.2.5 Calculator Keystrokes

We can use calculators for many things in math class.




The following instructions give the keystrokes for the TI-83/84 calculator, TI30 scientific calculator, and Casio fx-300 scientific calculator. To convert a fraction to a decimal, or a decimal to a fraction,

<u>TI-83 or TI-84</u>

Change a decimal to a fraction.

- ◇ Click on MATH key
- ◇ 1: Frac
- ◇ Enter

Change a fraction to a decimal.

- ◇ Click on MATH key
- ◇ 1: Dec
- ◇ Enter

<u>TI-30 (or a similar scientific calculator)</u>

Change a decimal to a fraction of a fraction to a decimal.

- ◇ Click on 2$^{\text{nd}}$key
- ◇ Click on PRB key
- ◇ Enter

<u>Casio fx-300 (or similar scientific calculator)</u>

Change a decimal to a fraction of a fraction to a decimal.

- ◇ Click on the S $\Leftrightarrow$ D button

### 1.2.6 Exercises 1.2

**Problem 1.1** Convert the percents to fractions and decimals.

(a) 17%

(b) 18.4%

(c) 135%

(d) 7.8%

(e) 1.26%

(f) 13.55%

(g) $1\frac{8}{9}\%$

(h) $179\frac{1}{8}\%$

(i) $10\frac{2}{3}\%$

**Solutions:**

(a) $\frac{17}{100}$, 0.17

(b) $\frac{23}{125}$, 0.184

(c) $\frac{27}{20}$, 1.35

(d) $\frac{39}{500}$, 0.078

(e) $\frac{63}{5000}$, 0.126

(f) $\frac{271}{2000}$, 0.1355

(g) $\frac{19}{900}$, 0.0189

(h) $\frac{1433}{800}$, 1.79125

(i) $\frac{8}{75}$, 0.1067

**Problem 1.2** Convert each decimal to a percent.

(a) 0.03 (b) 8 (c) 2.254 (d) 0.009

**Solutions:**

(a) 3% (b) 800% (c) 225.4% (d) 0.9%

**Problem 1.3** Convert each fraction to a percent.

(a) $\frac{5}{8}$ (b) $6\frac{4}{5}$ (c) $\frac{9}{8}$ (d) $3\frac{11}{12}$

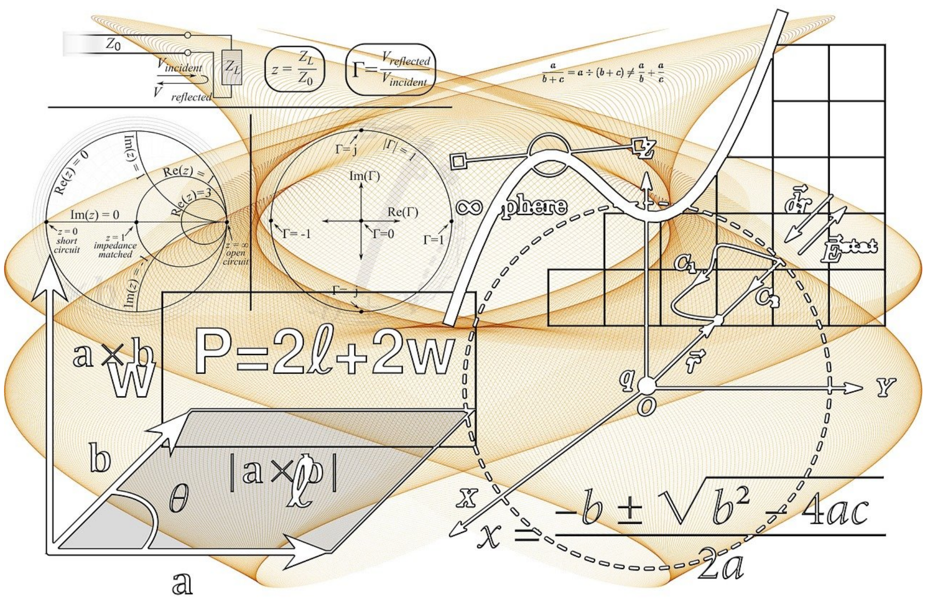
**Solutions:**

(a) 62.5% (b) 680% (c) 112.5% (d) 391.67%

## 1.3 Solving Equations in One Variable

These are some examples of solving equations in one variable. We will be using this information when we solve application problems involving percent, financial models, or anything that consists in solving for an unknown number.

When you see a variable by itself, it is understood to have a "1" in front of it. For example, if the variable is x, it is understood to be "1x". So, when you are adding x + 3x, you would get 4x since you are adding 1x + 3x.




■ **Example 1.14** Here are some examples of solving one-step equations.

<u>Adding or Subtracting a Number to Both Sides of an Equation</u>

$$
\begin{aligned}
x+5 &= 27 \quad \text{Original Equation} \\
x &= 27 \quad \text{Subtract 5 from both sides, this is your answer}
\end{aligned}
\tag{1.1}
$$

$$
\begin{aligned}
x-5 &= 27 \quad \text{Original Equation} \\
x &= 32 \quad \text{Add 5 to both sides, this is your answer}
\end{aligned}
\tag{1.2}
$$

<u>Multiplying by the Reciprocal to get a "1" on the variable</u>

$$\begin{aligned} 8x &= 27 \quad \text{Original Equation} \\ x &= \frac{27}{8} \quad \text{Subtract 5 from both sides, this is your answer} \end{aligned} \tag{1.3}$$

■

■ **Example 1.15** Here are some examples of solving two-step equations.

$$\begin{aligned} 8x+5 &= 27 \quad \text{Original Equation} \\ 8x &= 22 \quad \text{Subtract 5 from both sides} \\ x &= \frac{22}{8} \quad \text{Multiply by } \frac{1}{8} \text{ or divide by 8 on both sides, this is your answer} \end{aligned} \tag{1.4}$$

$$\begin{aligned} 8x-5 &= 27 \quad \text{Original Equation} \\ 8x &= 32 \quad \text{Add 5 to both sides} \\ x &= \frac{32}{8} \quad \text{Multiply by } \frac{1}{8} \text{ or divide by 8 on both sides} \\ x &= 4 \quad \text{Simplify, this is your answer} \end{aligned} \tag{1.5}$$

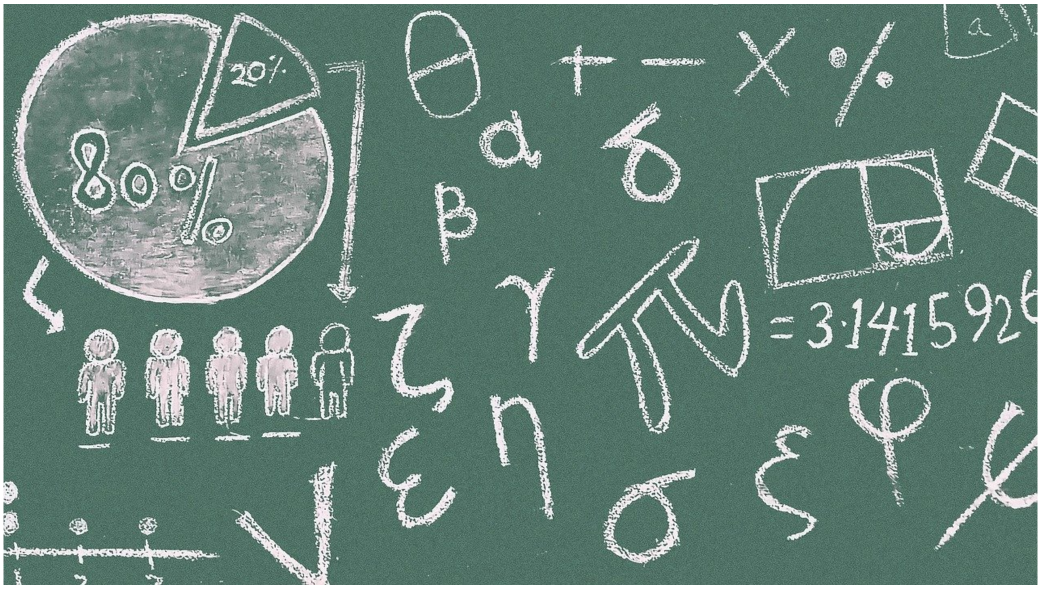
$$\begin{aligned} x+0.02x &= 5 \quad \text{Original Equation} \\ 1.02x &= 5 \quad \text{Add the 1x and 0.02x to get 1.02x} \\ x &= \frac{5}{1.02} \quad \text{Multiply by } \frac{1}{1.02} \text{ or divide by 1.02 on both sides} \\ x &= 4.9 \quad \text{Simplify, this is your answer} \end{aligned} \tag{1.6}$$

$$\begin{aligned} x-0.02x &= 5 \quad \text{Original Equation} \\ 0.98x &= 5 \quad \text{Subtract the 0.02x from 1x to get 0.98x} \\ x &= \frac{5}{0.98} \quad \text{Multiply by } \frac{1}{0.98} \text{ or divide by 0.98 on both sides} \\ x &= 5.1 \quad \text{Simplify, this is your answer} \end{aligned} \tag{1.7}$$

$$\begin{aligned} 3x-2 &= 5-x \quad \text{Original Equation} \\ 3x+x &= 5+2 \quad \text{x variables terms go to one side and number terms to other} \\ 4x &= 7 \quad \text{Multiply by } \frac{1}{4} \text{ or divide by 4 on both sides} \\ x &= \frac{7}{4} \quad \text{Simplify, this is your answer} \end{aligned} \tag{1.8}$$

■

### 1.3.1 Basic Application Problems




In solving application problems with percentages, there are three keywords that we need to know what they mean. Using these words and symbols, we can set up an equation and solve for the unknown. All percentages must be in decimal form to work the problems. The three keywords are:

- OF: means to multiply
- IS: means equals
- WHAT: means the unknown (any letter will work)

■ **Example 1.16** What is 55% of 60?

Pick out what we know in the problem.

- What: let's call that A
- Is: means equal
- 55% changes to 0.55
- Of: means to multiply

Now we have everything to set up our equation.

A = 0.55 * 60 = 33

■

■ **Example 1.17** Sherrie treated her parents to dinner at their favorite restaurant. The bill was \$73.25, and she wants to leave a 16% for the service. What would be the amount of the tip? What would be the total?


**Solution:** To figure out the tip, reword the problem as 16% of \$73.25 is the tip amount.

$0.16 * 73.26 = T$

$\$11.72 = T$

To find the amount for the dinner and tip, add the two totals together to get:

$\$73.25 + \$11.72 = \$84.97.$

Scherrie left \$84.97, which included a tip of \$11.72. ■

■ **Example 1.18** 36 is 75% of what number?

Pick out what we know in the problem.

- What: let's call that B
- Is: means equal
- 75% changes to 0.75
- Of: means to multiply

Now we have everything to set up our equation.

$36 = 0.75 * B$ or $B = \dfrac{36}{0.75} = 48$

■

■ **Example 1.19** What percent of 76 is 57?

Pick out what we know in the problem.

- What: let's call that C (and it wants a percent!)
- 75% changes to 0.75
- Is: means equal
- Of: means to multiply

Now we have everything to set up our equation.

C * 76 = 57

$$C = \frac{57}{76} = 0.75$$

Hence, we multiply by a hundred to obtain the percentage rate of $0.75 * 100 = 75$ or 75%. ■

■ **Example 1.20** One serving of oatmeal has 8 grams of fiber, which is 33% of the recommended daily amount. What is the total recommended daily amount of fiber?

**Solution:** To figure out the total recommended daily amount of fiber, reword the problem.

8 is 33% of the recommended daily amount of fiber.

$$F = \frac{8}{0.33} = 24.24 \text{ grams}$$

The total recommended daily amount of fiber is 24.24 grams.

■

■ **Example 1.21** Emma gets paid $3000 per month. She pays $750 a month for rent. What percent of her monthly pay goes to rent?

**Solution:** To figure out the total percent that goes to rent, reword the problem. What rental percent of Emma's paycheck is her rental payment?

$$3,000 * R = 750$$

$$R = \frac{750}{3000} = 0.25 * 100 = 25$$

Hence Emma pays 25% of her monthly paycheck towards her rent.

■

### Quick Practice - Basic Application Problems

**Exercise 1.3.1** Solve each of the problems.

(a) What number is 22.4% of 125?

(b) 37.5% of what number is 57?

(c) 7.2 is what percent of 16?

(d) What number is 87.5% of 70?

(e) 5.6 is what percent of 40?

(f) Aaron notes that 15% of his class is absent. If the class has 80 students, how many students are absent?

(g) Misty answered 19 of 25 possible questions on her her biology test correctly. What percent of the questions did she mark correctly?

(h) A family has traveled 102 miles of a planned trip. This is 23% of the total distance they must travel on the trip. Find to the nearest mile, the total distance they will travel on their trip.

(i) Towana took her family out to dinner for a birthday celebration. The bill was \$165.35. The left a \$35.65 tip. What was the percent of tip that Towana left?

**Solutions:**

(a) 28
(b) 152
(c) 45%
(d) 61.25
(e) 14%
(f) 12
(g) 76%
(h) 444 miles
(i) 22%

### 1.3.2 Commissions, Sales Tax, Discounts, and Mark-Ups

Salespeople often receive a commission, which is a percent of their sales, for their income. This may be their only income, or they may earn a commission in addition to their hourly wages or salary. The commission they make is calculated as a certain percent of the price of each item they sell. That percent is called the rate of commission.

The types of professions that pay commissions include car sales, real estate brokers, travel agents, insurance sales, stockbrokers, basically anyone who deals in selling items or services.


**Definition 1.3.1** A **commission** is a percentage of total sales as determined by the rate of commission. It is calculated by (the percent must be in decimal form):

Percentage rate of commission * total sales = commission

■ **Example 1.22** Helene is a Realtor. She receives $3\%$ commission when she sells a house. How much commission will she receive for selling a home that costs $\$395,000$?

We can set up the problem using the commission formula. Let $C$ = the commission, 0.03 = the percentage rate, and \$395,000 be the sales

$$C = 0.03 * 395,000 = 11,850$$

Helene made \$11,850 from selling a house that costs \$395,000.

■

■ **Example 1.23** Homer received \$1,140 commission when he sold a car for \$28,500. What rate of the commission did he get?

We can set up the problem using the commission formula. Let the commission $C = \$1,140$, $P$ = the percentage rate, and $\$28,500$ be the sales

$$P * 28,500 = 1,140$$

$$P = \frac{1140}{28500} = 0.04$$

$$P = 0.04 * 100 = 4\%$$

Homer made a commission of 4% from selling a car that was worth \$28,500.

■

Whenever people purchase items such as food, clothes, gardening materials, or appliances, a certain percentage of taxes must be paid on these items. This sales tax amount varies depending on what county you shop in, and this money is used to help fund the state and local budgets for schools, roads, and fire departments. The companies collect this money and send it to the state.

**Definition 1.3.2** A **sales tax** is a percentage of tax that is paid on the price of purchased items. It can be calculated by (the percent must be in decimal form):

Cost of Item(s) + % of tax * Cost of Item(s) = Total Cost

The amount of tax is the % of tax * Cost of the Item(s) Purchased

■ **Example 1.24** Jennifer bought a new winter coat in St. Louis for \$250.00. The sales tax was 8.2%. Find the amount of the tax and the total cost.

We can set up the problem using the sales tax formula. Cost of Item = \$250.00, Total Cost = $T$, % of tax = 8.2%.

$$T = 250 + .082(250) = 270.5$$

Jennifer paid \$270.50 for her new coat.

Tax paid: 0.82 * 250 = 20.5 <u>OR</u> Tax paid: 270.5 - 250.00 = 20.5

Kim paid \$20.50 tax on her new coat.

(You can find the tax by multiplying the % of sales tax and the cost of the coat <u>OR</u> by taking the total cost of the coat and subtracting the cost of the coat. The answer will be the same.) ■

■ **Example 1.25** Diego bought a new smartphone for \$499 plus tax. He was surprised when he got the receipt and saw that the tax was \$42.42. What was the sales tax rate for this purchase?

We can set up the problem using the sales tax formula.

Cost of Item = \$499 - \$42.42 = \$456.58, % of tax = use P, Total Cost = \$499

$456.58 + P(456.58) = 499$

$P(456.58) = 499 - 456.58$

$(456.58) = 42.42$

$P = \dfrac{42.42}{456.58}$

$P = 0.0929 * 100 = 9.29\%$ (Since this is asking for a percentage rate, you must multiply by 100.)

Diego paid 9.29% sales tax on his new smartphone.

■

■ **Example 1.26** You purchased a new Ipad in Atlanta, which has an 8.9% sales tax. The total cost (including tax) for your new iPad was \$505.40. What was the pre-tax price?

We can set up the problem using the sales tax formula, where $I$ is the cost of the item, The percentage % of tax is 8.9%, and the Total Cost is \$505.40.

$I + I(.089) = 505.4$

$1.089I = 505.4$

$I = \dfrac{505.4}{1.089}$

$I = \$464.10$

The Ipad you purchased cost \$464.10 before tax.

■

Everyone loves a good deal, especially when they can save some money. These are called discounts, and that's why sellers have sales and offer coupon deals. Sometimes you can get a specific dollar amount off of your purchase, or you can get a percent off of your purchase.

It would help if you looked closely at the signs since clearance racks sometimes say, "Take an additional 25% off the last reduced price", which doesn't mean the original price. Sales help retailers reduce their inventory to bring new stock in for their customers.

Mark-ups are very common in retail settings. The price a retailer pays for an item is called the wholesale price, and then the retailer adds a mark-up to the wholesale price to get the list price, which is the price he sells the item for. The mark-up is usually calculated as a percent of the wholesale price. This is how the retailer makes money.

**Definition 1.3.3** A **discount** is a percentage that we reduce from the original price. It can be calculated by (the percent must be in decimal form):

Original Price of Item(s) - % of discount * Original Price of Item(s) = SALE PRICE

The amount of the discount is the % of discount * Cost of the Item(s) Purchased


**Definition 1.3.4** The **mark-up** is the amount added to the wholesale price. It can be calculated by (the percent must be in decimal form):

Wholesale Price of Item(s) + % of mark-up * Wholesale Price of Item(s) = LIST PRICE

The amount of the mark-up is the % of mark-up * Wholesale Price of the Item(s) Purchased

■ **Example 1.27** Oscar bought a barbecue grill that was discounted 65% from an original price of $395. Find the sale price and the amount of the discount.

We can set up the problem using the discount formula, where the original price of the item = $395, the % of discount = 65%, and the Sale Price = $G$, using all this we obtain:

$395 - 0.65(395) = G$

$138.25 = G$

The sale price for the barbecue grill was $138.25.

Amount of discount: 0.65 * 395 = 256.75

OR Amount of discount: 395 - 138.25 = 256.75

The amount of discount of Oscar's barbecue grill was $256.75.

(You can find the discount amount by multiplying the % of the discount, and the original price of the barbecue grill OR by taking the original price of the barbecue grill and subtracting the sale price of the barbecue grill. The answer will be the same.)

■

■ **Example 1.28** Carmax bought Pablo's Toyota for $8,500. The car resale store marked the price up 35%. Find the list price and the amount of the mark-up.

We can set up the problem using the mark-up formula as follows. First, set Wholesale Price of Item = $8,500, next we establish the % of mark-up is 35%, finally List Price is $T$

$8,500 + .35(8,500) = T$

$11,475 = T$

The Auto Resale Store has the list price for Pablo's Toyota at $11,475.00.

Amount of mark-up: 0.35 * 8,500 = 2,975 OR Amount of mark-up: 11,475 - 8,500 = 2,975

The amount of mark-up on Pablo's Toyota is $2,975.00.

(You can find the amount of mark-up by multiplying the % of the mark-up, and the wholesale price of the Toyota OR by taking the list price of the Toyota and subtracting the wholesale price of the Toyota. The answer will be the same.) ■

■ **Example 1.29** Norma paid \$150 for a pair of boots initially priced at \$620. What percent of the original price did Norma pay?

Determine if this is a discount or mark-up problem. Since the original price is higher than what she paid, it is a discount problem. Use the discount formula.

Original Price of Item = \$620, % of discount = W, SALE PRICE = \$150

$620 - W(620) = 150$

$-W(620) = -470$

$-W = \dfrac{-470}{620}$

$-W = -0.7581$ (Divide by a -1)

$W = 0.7581 = 75.81/100 = 75.81\%$

Norma's boots were discounted 75.8% from the original cost.

■

■ **Example 1.30** Tom bought a pair of jeans for \$63.00; the price paid included a discount of 45%. What was the original price of a pair of jeans?

We can set up the problem using the discount formula.

Original Price of Item = J, % of mark-up = 45%, SALE PRICE = \$63

J - J (0.45) = 63

0.55J = 63

J= $\dfrac{63}{0.55} = 114.55$

The original price of Tom's jeans were \$114.55.

■

Once you get an answer, you can substitute all the numbers into the equations to see if you get the same answers. For example, on the above problem with Norma's jeans, you could use this to check:

$$\$114.55 - (0.45)(\$114.55) = \$63.0025.$$

This answer is close enough that it can be considered as correct.

**Quick Practice - Commissions, Sales Tax, Discounts, Mark-Ups**

**Exercise 1.3.2** Solve each of the problems.

(a) Bill earns a commission on all sales he makes. He sells a bed for \$591 and earns a commission of \$43. Find the percent commission, rounded to the nearest tenth of a percent.

(b) Alan bought a used bicycle for \$115. After re-conditioning it, he added 225% mark-up and then advertised it for sale. What is the list price and the amount of mark-up?

(c) Hector receives 17.5% commission when he sells an insurance policy. How much commission will he receive for selling a policy for \$4,910?

(d) Sylvia bought a new microwave and the cost was \$138.68, including tax. The rate of tax is 7.5%. What was the price of the microwave before tax?

(e) Sophie saw a dress she liked on sale for \$15 off. The original price of the dress was \$96. What was the percentage of the discount on the dress?

(f) Shawna bought a mixer for \$300. The sales tax on the purchase was \$19.50. What is the sales tax rate?

(g) Kathy wants to buy a camera that lists for \$389. The camera is on sale with a 33% discount. Find the sale price and the amount of money she saved.

**Solutions:**

(a) 7.3%

(b) \$373.75, \$258.75

(c) \$859.25

(d) \$129.00

(e) 16%

(f) 6.5%

(g) \$260.63, \$128.37

(h) 85%

■

### 1.3.3 Simple Interest

Interest is a fee or a charge for borrowing money, typically a percent rate charged per year. When a person takes out a loan, most lenders charge interest on the loan. We can compute simple interest by finding the interest rate percentage of the amount borrowed, then multiplying it by the number of years interest is earned.

Bank customers can earn interest on money when they deposit money into a bank account. The bank will pay customers interest and a specific percentage rate on your principal. The principal is the money that has been deposited into the bank account.

**Definition 1.3.5** **Simple Interest** is calculated by the formula

$$I = P * r * t,$$

where

- I = Amount of interest earned
- P = Principal, The amount of money borrowed or deposited in the bank.
- r = percentage rate, written as a decimal
- t = time, in years

Working on application problems with interest is like working on the previous problems in this chapter. Identify the numbers that go with the letters $I, P, r$, and $t$, substitute them into the formula, and solve.

■ **Example 1.31** Areli invested a principal of $950 in her bank account with an interest rate of 3%. How much interest did she earn in 5 years?

Interest = I, Principal = $950, r = 0.03, t = 5

$I = 950 * 0.03 * 5$

$I = \$142.50$

Areli will earn $142.50 interest on her bank account in 5 years.

■

■ **Example 1.32** Eduardo noticed that his new car loan papers stated that with an interest rate of 7.5%, he would pay $6,596.25 in interest over five years. How much did he borrow to pay for his car?

Interest = $6,596.25, Principal = P, r = 0.075, t = 5

$6,596.25 = P * 0.075 * 5$

$6,596.25 = 0.375P$

$\frac{6596.25}{0.375} = P$

$\$17,590.00 = P$

Eduardo will borrow $17,590.00 to buy a new car.

■

■ **Example 1.33** Loren lent his brother $3,000 to help him buy a car. In 4 years, his brother paid him back the $3,000 plus $660 in interest. What was the rate of interest?

Interest = $660, Principal = $3,000, r = r, t = 4

660 = 3,000 * r * 4

660 = 12,000r

$r = \frac{660}{12,000} = 0.055 = 5.5/100 = 5.5\%$

The rate of interest that Loren's brother paid back was 5.5%.

■

■ **Example 1.34** Robin deposited \$31,000 in a bank account with an interest rate of 5.2%. After so many years, she withdrew her money, and it had earned \$4,836 in interest. How many years had Robin kept her money in the bank?

Interest = \$4,836, Principal = \$31,000, r = 0.052, t = t

4,836 = 31,000 * 0.052 * t

4,836 = 1,612t

$t = \frac{4832}{1612} = 2.997$

Robin kept her money in the bank for 3 years.

■

■ **Example 1.35** Areli borrowed \$3,900 from her parents for the down payment on a car and promised to pay them back in 15 months at a 4% rate of interest. How much interest did she owe her parents?

Interest = I, Principal = \$3,900, r = .04, t = $\frac{15}{12}$ = 1.25

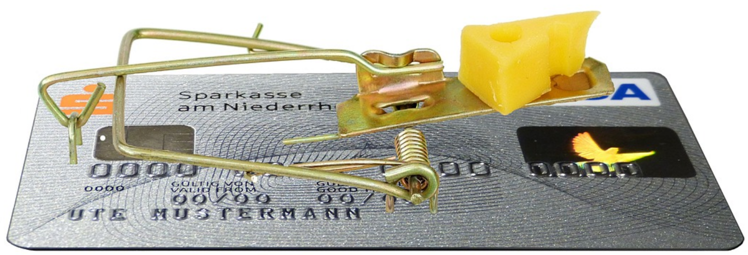
The time must be in years but since this is in months, just convert it into years.

$I = 3,900 * 0.04 * 1.25 = \$195.00$

Areli will owe her parents \$195 in interest in 15 months.

■

**Quick Practice - Simple Interest**




**Exercise 1.3.3** Solve each of the problems.

(a) Hilaria borrowed \$8,000 from her grandfather to pay for college. Five years later, she paid him back the \$8,000, plus \$1,200 interest. What was the rate of interest?

(b) Caitlin invested \$8,200 in an 18-month certificate of deposit paying 2.7% interest. How much interest did she earn from this investment?

(c) Terrence deposited \$5,720 in a bank account with an interest rate of 6%. How much interest was earned in 4 years?

(d) Joshua's computer loan statement said he would pay \$1,244.34 in interest for a three-year loan at 12.4%. How much did Joshua borrow to buy the computer?

**Solutions:**

(a) 3% (b) \$332.10 (c) \$1,372.80 (d) \$3,345.00

### 1.3.4 Absolute and Relative Change

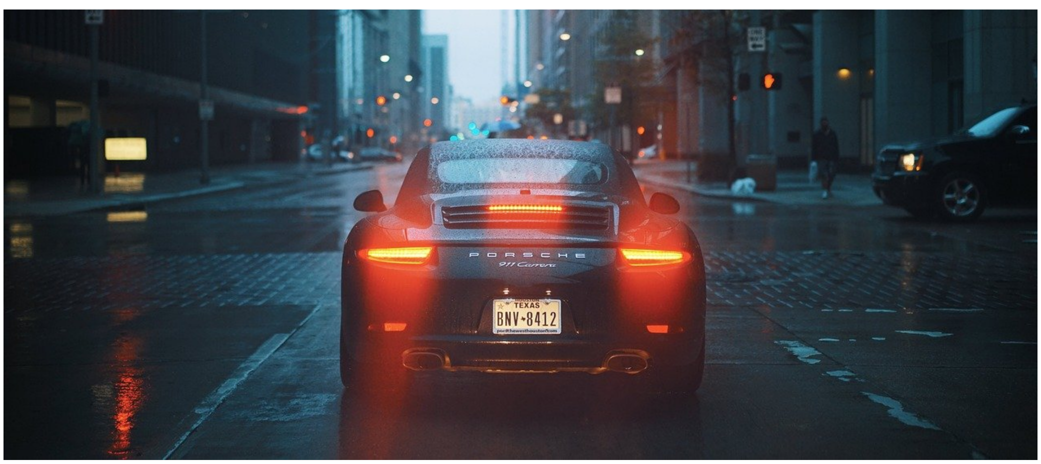
Frequently, percentages are used to compare how a quantity has changed over time. We use absolute and relative change when making this type of comparison.

**Definition 1.3.6** To find the **Absolute** and **Relative** Change, you must be given two quantities.

- Absolute change = ending quantity – starting quantity
- Absolute change has the **same units** as the original quantity.
- Relative change = $\dfrac{\text{absolute change}}{\text{starting quantity}} * 100$
- Relative change gives a **percent** change.

■ **Example 1.36** Find the absolute and relative change.

The value of a car dropped from \$7,400 to \$6,800 over the last year. What percent decrease is this?




The absolute change is going to be the dollar value change, which was negative.

This difference is: 6,800 - 7,400 = -600 (Absolute Change)

Since we are computing the decrease relative to the starting value, we compute this percent out of $\$7,400$.

$$\frac{-600}{7400} = -0.081 * 100 = -8.1\% \text{ (Relative Change)}$$

■

**Quick Practice - Absolute and Relative Change**


Figure 1.4: Data for the exercise from: `tinyurl.com/y2xwms2d`

**Exercise 1.3.4** The U.S. federal debt at the end of 2001 was \$5.77 trillion and grew to \$6.20 trillion by the end of 2002. At the end of 2005, it was \$7.91 trillion and grew to \$8.45 trillion by the end of 2006[1].

(a) Calculate the absolute and relative increase for 2001 to 2002 and 2005 to 2006.

(b) Which year saw a larger increase in federal debt?

**Solutions:**

(a) Absolute Change 2001 to 2002 = 0.43 trillion

Relative Change 2001 to 2002 = 7.45%

Absolute Change 2005 to 2006 = 0.54 trillion

Relative Change 2005 to 2006 = 6.83%

(b) 2005 to 2006 saw a larger absolute increase, but a smaller relative increase.

### 1.3.5 Exercises 1.3

**Problem 1.4** Solve each problem.

(a) What number is 45% of 120?

(b) 81 is 75% of what number?

(c) What percent of 1,500 is 540?

(d) 840 is what percent of 480?

(e) 11.5% of what number is 108.1?

(f) What number is 24% of 112?

**Solutions:**

(a) 54

(b) 108

(c) 35%

(d) 175%

(e) 940

(f) 26.88

**Problem 1.5** Solve each problem.

(a) Alison was late paying her credit card bill of $249. She was charged a 5% late fee. What was the amount of the late fee?

(b) The nutrition fact sheet at a fast-food restaurant says the fish sandwich has 380 calories, and 171 calories are from fat. What percent of the total calories is from fat?

(c) Cherise deposits 8% of each paycheck into her retirement account. Her last paycheck was $1,485. How much did Cherise deposit into her retirement account?

(d) Douglas gets paid $3,200 per month. She pays $960 a month for rent. What percent of her monthly pay goes to rent?

(e) A lender requires a minimum down payment of 12% of the home's value. You have $22,020 cash available to use as a down payment toward a home. Determine the maximum home value that you can finance.

(f) Dontay is a realtor and earned $11,250 commission on the sale of a $375,000 house. What is his rate of commission?

(g) Farrah works in a jewelry store and receives 12% commission when she makes a sale. How much commission will she receive for selling a $8,125 ring?

(h) As a waitress, Emily earned $420 in tips on sales of $2,625 last Saturday night. What was her rate of commission?

(i) Shawna bought a mixer for $300. The sales tax on the purchase was $19.50.

(j) The cost of a 6-drawer dresser $1,199. The sales tax rate is 5.125% of the purchase price. Find the sales tax amount and the total cost amount.

(k) The list price for a candle at a store is $18.95. The store has a mark-up of 85%. Find the wholesale cost of the candle.

(l) Marcus bought a bike on sale for $175.00 when the original price was $235.00. What was the percent of discount?

(m) Jackie buys raincoats for her store. She purchased them at wholesale for $65.00 and marked them up to $70.20. What is the percentage of the mark-up?

(n) The balance on a 6-year loan is $10,222. If the principal borrowed was $7,600, what was the simple interest rate (as a percent)?

(o) How much simple interest is earned if $5,800 is invested at 3.25% per year for 21 months?

(p) Find the principal invested if $15,222.57 of interest was earned in 6 years at a 10.28% rate.

(q) Kara deposited $14,000.00 in the bank with an interest rate of 4.5%. At the time she withdrew her money, she earned $4725.00. How many years did she wait to withdraw her money?

(r) The average amount of student loan debt in 2010 was $25,250, and in 2021 was $37,172. Find the absolute and relative change.

**Problem 1.6** A new tractor costs \$1,768.55, including a sales tax of 7.25%. How much was the sales tax amount?


**Problem 1.7** Tim wanted a new game for his x-box, and he found one on sale for \$57.80. The discount was 12%. Find the original price of the game.

**Solutions:**

(a) $12.45

(b) 45%

(c) $118.80

(d) 30%

(e) $183,500.00

(f) 3%

(g) $975.00

(h) 16%

(i) 6.5%

(j) $61.45, $1,260.45

(k) $10.24

(l) 25.6%

(m) 8%

(n) 5.75%

(o) $329.88

(p) $24,679.91

(q) 7.5 years

(r) $11,922 and 47.2%

(1.6) $119.50

(1.7) $65.68

## 1.4 Ratios and Proportions

Ratios are used to make comparisons, so it's not surprising that they are used heavily in the finance, design, and science fields.

Figure 1.5: Ratios are used to compare and produce proportions of things.

People looking to invest or purchase a company will analyze specific ratios to determine the health of a company to make an informed decision to move forward with the investment or purchase. Loan companies will look at prospective client debt to income ratios to determine the client's risk of paying back the loan.

People who work in fields where blueprints are used, such as building or commercial architects, landscaping architects, and interior design architects, work with ratios. It is impossible to draw large designs to actual size, so a ratio represents the sizes. For example, one inch may equal one foot on the blueprint.

At a compound pharmacy, a pharmacist can make specific medications when unavailable. Therefore, ratios and proportions are needed to mix the correct amount of medicine.

These are a few samples of how we might use a ratio.

1. To describe the cost of a month's rent compared to the income earned in one month.
2. To compare the number of elephants to the total number of animals in a zoo.
3. To compare the number of calories per serving in two different brands of ice cream.

### 1.4.1 Ratios

First, let's introduce the notion of ratio:

**Definition 1.4.1** A **ratio** is used to compare amounts or quantities or describe a relationship between two parts or portions that are measured in the **same unit**.

The ratio of a to b is written as:

- a to b
- $\frac{a}{b}$ (simplified)
- a : b

You do not have to write the ratio so that the lesser quantity comes first; the important thing is to keep the **relationship consistent**.

■ **Example 1.37** Paul compares the number of calories in personal pizza to the calories in a sub sandwich. The pizza has 600 calories, while the sandwich has 450 calories.

1. Write a ratio that represents the number of calories in the pizza compared to the calories in the sub sandwich.

   The relationship is $\frac{\text{calories in pizza}}{\text{calories in sub sandwich}} = \frac{600}{450} = \frac{4}{3}$

   The ratio of calories in the pizza to calories in the sub sandwich is $\frac{4}{3}$, 4:3, or 4 to 3.

2. Write a ratio that represents the amount of calories in the sub sandwich compared to the calories in the pizza.

   The relationship is $\frac{\text{calories in sub sandwich}}{\text{calories in pizza}} = \frac{450}{600} = \frac{3}{4}$

   The ratio of calories in the sub sandwich to calories in the pizza is $\frac{3}{4}$, 3:4, or 3 to 4.

■

The above example showed that even though the numbers we used were the same, the ratio depended on the relationship needed. How would this differ if the question asked for the proportion representing the number of calories in the pizza compared to the total number of calories in the pizza and sub sandwich?

■ **Example 1.38** Paul compares the number of calories in personal pizza to the calories in a sub sandwich. The pizza has 600 calories, while the sandwich has 450 calories.

Write a ratio representing the number of calories in the pizza compared to the total number of calories in the pizza and sub sandwich.

The relationship is $\dfrac{\text{calories in pizza}}{\text{total calories}} = \dfrac{600}{1050} = \dfrac{4}{7}$

The ratio of calories in the pizza to total calories is $\dfrac{4}{7}$, 4:7, or 4 to 7.

■

The numbers may be given as decimals or in fraction form when writing a ratio. By using the calculator, we can make the conversions quickly and easily. (Refer to section 1.2 on how to change from a decimal to a fraction on your calculator.)

■ **Example 1.39** Write each ratio as a fraction in simplified form.

(a) 3.4 to 15.3 $\rightarrow \dfrac{3.4}{15.3} = \dfrac{2}{9}$

(b) 2.7 to 0.54 $\rightarrow \dfrac{2.7}{0.54} = \dfrac{5}{1}$ (The 1 must remain in the denominator.)

(c) $1\frac{1}{4}$ to $2\frac{3}{8} \rightarrow \dfrac{1.25}{2.375} = \dfrac{10}{19}$ (Could have changed into improper fraction, if desired.)

(d) $1\frac{1}{4}$ to $2\frac{3}{8} \rightarrow \dfrac{1.25}{2.375} = \dfrac{10}{19}$ (Could have changed into improper fraction, if desired.)

(e) $6\frac{5}{6}$ to $3\frac{1}{9}$

$$6\frac{5}{6} = \frac{\frac{41}{6}}{\frac{28}{9}}$$
$$= \frac{41}{6} \div \frac{28}{9}$$
$$= \frac{41}{6} * \frac{9}{28} = \frac{123}{56}$$

Since the fractions are repeating, we must convert them into improper fractions.

■

**Quick Practice - Ratios**

**Exercise 1.4.1** Write each of the ratios as a fraction. Leave in reduced fraction form.

1. 42 to 48
2. \$1.21 to \$0.44
3. $1\frac{2}{3}$ to $2\frac{5}{6}$
4. 27 inches to 1 foot (there are 12 inches in one foot)
5. Luisa invites a group of friends to a party. Including Luisa, there are 22 people, 10 of whom are women. Find the ratio of women to men. Find the ratio of women to the total number of people attending the party.

**Solutions:**

1. $\frac{7}{8}$ 2. $\frac{11}{4}$ 3. $\frac{10}{17}$ 4. $\frac{9}{4}$ 5. $\frac{5}{6}, \frac{5}{11}$

### 1.4.2 Rates and Unit Rates

A rate is a comparison that provides information such as dollars per hour, feet per second, miles per hour, and dollars per quart. The word "per" usually indicates you are dealing with a rate.

When writing a fraction as a rate, we write the numerator and the denominator, both with their units. When rates are simplified, the units remain in the numerator and denominator.

**Definition 1.4.2** A **rate** is a ratio that compares two different quantities that have other units of measure. They can be written using words, a colon, or a fraction.

Susan can read 92 pages read in 8 minutes

92 pages read per 8 minutes

92 pages read : 8 minutes

$$\frac{92 \text{ pages read}}{8 \text{ minutes}} = \frac{23 \text{ pages read}}{2 \text{ minutes}}$$

Although this rate is valid, it doesn't help us. We would rather know how many pages Susan read in 1 minute.

**Definition 1.4.3** A **unit rate** compares a quantity to one unit of measure. It is written with a denominator of one unit.

We know Susan can read 23 pages in 2 minutes; we can divide to find how many pages she can read in one minute.

$$\frac{11.5 \text{ pages read}}{1 \text{ minute}} = 11.5 \text{ pages read/minute}$$

In a previous we talked about discounts, we know that everyone likes to get a good deal when shopping. It is advantageous to compare units before making your purchase.

Most grocery stores list the unit price of each item on the shelves. For example, a store may offer canned corn at \$0.50 per can or six cans of corn for \$2.50. That would give you a savings of \$0.50 if you bought six cans of corn. Only you can decide if you want or need six cans of corn, but at least you can make an informed decision.

**Definition 1.4.4** A **unit price** is a unit rate that gives the price of one item. The better buy is the item with the lower unit price.

■ **Example 1.40** The grocery store charges \$3.99 for a case of 24 bottles of water. What is the unit price?


We are asked to find the unit price, which is the price per bottle, so the ratio should be set up with $\frac{\text{price}}{\text{bottle}}$

$$\frac{\$3.99}{24 \text{ bottles}} = \frac{\$0.16625}{1 \text{ bottle}} = \frac{\$0.17}{1 \text{ bottle}}$$

Each bottle of water costs about \$0.17. ■

■ **Example 1.41** Paul is shopping for laundry detergent. At the grocery store, the liquid detergent is priced at \$14.99 for 64 loads of laundry, and the same brand of powder detergent is priced at \$15.99 for 80 loads. Which is the better buy, the liquid or the powder detergent?

We want to compare the cost per load, so the ratio should be set up with $\frac{\text{price}}{\text{load}}$

1. Liquid Detergent: $\frac{\$14.99}{64 \text{ load}} = \frac{\$0.23421875}{1 \text{ load}} = \frac{\$0.23}{1 \text{ load}}$

   The cost of the liquid detergent is about \$0.23 a load.

2. Powder Detergent: $\frac{\$15.99}{80 \text{ load}} = \frac{\$0.199875}{1 \text{ load}} = \frac{\$0.20}{1 \text{ load}}$

   The cost of the liquid detergent is about \$0.20 a load.

   The cost of the powder detergent is a better buy.

■

**Quick Practice - Rates and Unit Rates**

**Exercise 1.4.2** Find the unit rates or prices and answer the questions.

1. You need tomato sauce for your enchiladas. You can buy three cans of tomato sauce for \$1.49 or 5 cans of tomato sauce for \$2.39.Which is the best buy?

2. Jeff's automobile travels 422 miles on 15 gallons of gasoline. Tony's automobile travels 354 miles on 13 gallons of gasoline. Which automobile gets the better gas mileage?

3. Frannie works 5.5 hours and gets paid \$95.55. Anna works 6.75 hours and gets paid \$117.31. Who gets paid more per hour?

**Solutions:**

1. 3 cans = \$0.50/can, 5 cans = \$0.48/can; The cost of the 5 cans is the best buy.

2. Jeff = 28.1/mpg, Tony = 27.2/mpg; Jeff's car gets the better gas mileage.

3. Frannie = \$17.37, Anna = \$17.38; Anna gets paid more per hour.

### 1.4.3 Proportions

In the last section, we learned that a ratio compares two quantities. If we take two ratios and determine that they are equal, they are said to be a proportion. If you know one ratio in a proportion, you can use that information to find values in the other equivalent ratio. Using proportions can help you solve problems by increasing a recipe to feed a larger crowd of people, creating a design with certain consistent features, or enlarging or reducing an image to scale.

You can set up a proportion to determine the length of the enlarged photo. For example, imagine you want to enlarge a 5-inch by 8-inch photograph to fit a wood frame that you purchased. If you wish the shorter edge of the enlarged-photo to measure 10 inches, how long does the picture have to be for the image to scale correctly?

**Definition 1.4.5** A **proportion** is an equation in the form:

$$\frac{a}{b} = \frac{c}{d}$$

where $b \neq 0$ and $d \neq 0$.

- The proportion states two ratios or rates are equal.
- The proportion is read "a is to b, as c is to d".

The ratios need to be set up consistently, either vertically or horizontally, to determine if the proportion is equal.

**Example 1.42** Juanita has two different-sized containers of lemonade mix. She wants to compare them to see if they are the same. The first container has 40 ounces and serves ten people, and the second container has 84 ounces and serves 21 people.

1. This ratio is set up vertically with the ounces in the numerator and the serving sizes in the denominator.

$$\frac{40 \text{ ounces}}{10 \text{ servings}} = \frac{84 \text{ ounces}}{21 \text{ servings}} \rightarrow \frac{4 \text{ ounces}}{1 \text{ servings}} = \frac{4 \text{ ounces}}{1 \text{ servings}}$$

   Since the ratios are the same, they are a proportion.

2. This ratio is set up horizontally with the ounces in the first ratio and the serving sizes in the second ratio. In this setup, the numerators must come from the same scenario.

$$\frac{40 \text{ ounces}}{84 \text{ ounces}} = \frac{10 \text{ servings}}{21 \text{ servings}} \rightarrow \frac{40}{84} = \frac{10}{21} \rightarrow \frac{10}{21} = \frac{10}{21}$$

   The units cancel because they are the same in the numerator and denominator. Since the ratios are the same, they are a proportion.

■

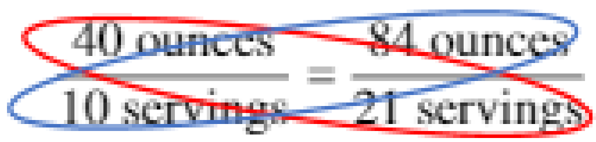
Another way to determine if the ratios are equal is to find the cross-product.

**Definition 1.4.6** For any proportion in the form:

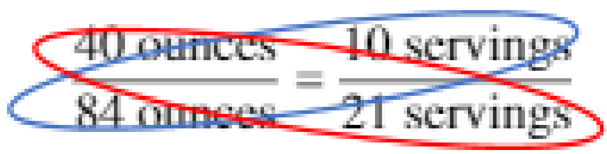
$$\frac{a}{b} = \frac{c}{d} \text{ where } b \neq 0 \text{ and } d \neq 0.$$

the **cross products** are equal if a * d = b * c.

In the example with Juanita and the lemonade mix, we could have used cross-products to determine if the ratios were equal.




40 * 21 = 840 and 10 * 84 = 840




40 * 21 = 840 and 84 * 10 = 840

Cross product function as a natural way to validate our proportions; of course, sometimes elements of the ratios we want to find are not going to be present, which leads to the necessary presence of equations in our proportions.

### 1.4.4 Solving Equations using Proportions

If you know that the relationship between quantities is proportional, you can use proportions to find the missing amounts. Use the cross product and solve for the unknown variable.

■ **Example 1.43** Solve the proportions.

1. $\frac{x}{63} = \frac{4}{7}$

   To solve this proportion we cross multiply and obtain

   $$7x = 252 \to x = 36$$

2. $\frac{2.7}{m} = \frac{0.9}{0.2}$

   To solve this proportion we cross multiply and obtain

   $$0.54 = 0.9m \to m = 0.6$$

■

■ **Example 1.44** Set-up the ratios and solve the proportions.

1. Pediatricians sometimes prescribe acetaminophen to children. They prescribe 5 milliliters (ml) of acetaminophen for every 25 pounds of the child's weight. If Zoe weighs 80 pounds, how many milliliters of acetaminophen will her doctor prescribe?

   $$\frac{5 \text{ ml}}{25 \text{ pounds}} = \frac{\text{Z ml}}{80 \text{ pounds}}$$
   $$400 = 25\text{Z}$$
   $$16 = \text{Z}$$

   Zoe needs 16 ml of acetaminophen.

2. Yaneli loves Starburst candies but wants to keep her snacks to 100 calories. If the sweets have 160 calories for eight pieces, how many can she have in her snack?

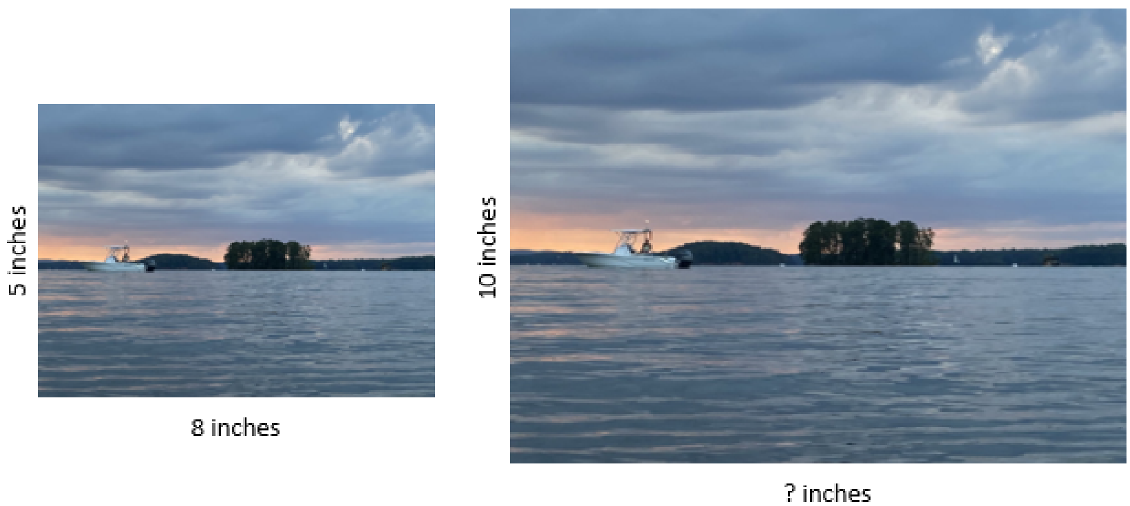
   $$\frac{160 \text{ calories}}{8 \text{ pieces}} = \frac{100 \text{ calories}}{\text{C pieces}}$$
   $$160\text{C} = 800$$
   $$\text{C} = 5$$

   Yaneli can have 5 pieces of candy in her snack.

■

Let's go back to the example mentioned at the beginning of this section, where we talked about enlarging a photograph. Imagine you want to widen a 5-inch by 8-inch photograph to make the length 10 inches and keep the proportion of the width to length the same. You can set up a ratio to determine the width of the enlarged photo.

■ **Example 1.45** Find the length of a photograph whose width is 10 inches and whose proportions are the same as a 5-inch by 8-inch photograph.




$$\frac{5 \text{ wide}}{8 \text{ long}} = \frac{10 \text{ wide}}{\text{P long}}$$

$$5P = 80$$

$$P = 16$$

The photograph needs to be enlarged to a length of 16 inches.

■

**Quick Practice - Proportions**

**Exercise 1.4.3** Answer each of the questions.

(a) Solve the proportion: $\dfrac{5}{a} = \dfrac{65}{117}$

(b) Solve the proportion: $\dfrac{2.6}{3.9} = \dfrac{c}{3}$

(c) A recipe making two dozen cookies requires $1\frac{3}{4}$ cups of flour. How much flour will be needed to make 4 dozen cookies?

(d) If 3 dog bones cost \$1.23, what will be the cost for 13 dog bones?

**Solutions:**

(a) 9

(b) 2

(c) $3\frac{1}{2}$

(d) \$5.33

### 1.4.5 Unit Conversions

We can take proportion problems a step further by multiplying a quantity by rates to change the units, and this is called a unit conversion.

If the question was asked, "If your car can drive 300 miles on a tank of 15 gallons of gas, how far can it drive on 40 gallons of gas?" We could solve this using a proportion.

$$\frac{300 \text{ miles}}{15 \text{ gallon}} = \frac{x \text{ miles}}{400 \text{ gallons}}$$

$$15x = 12,000$$

$$x = 800$$

On 40 gallons of gas, you can travel 800 miles.

How could we solve the problem if the question was, "How many gallons of gas are needed to travel 50 miles?" This time, we would have to set up a unit conversion so that only gallons are left at the end when we cancel out our units.

$$\frac{50 \cancel{\text{mi}}}{1} * \frac{1 \text{ gal}}{20 \cancel{\text{mi}}} = \frac{50 \text{ gal}}{20} = 2.5 \text{ gallons}$$

You can travel 50 miles on 2.5 gallons of gas.

The conversion chart below will help with setting up your unit conversions.

| U.S. System Measurements | |
|---|---|
| **Time:** | **Volume:** |
| 1 min = 60 sec | 3 t = 1 T |
| 1 hr = 60 min | 16 T = 1 C |
| 1 day = 24 hr | 1 C = 8 fl oz |
| 1 wk = 7 days | 1 pt = 2 C |
| 1 yr = 365 days | 1 qt = 2 pt |
| | 1 gal = 4 qts |
| **Length:** | **Miscellaneous:** |
| 1 ft = 12 in | 1 yd$^2$ = 9 ft$^2$ |
| 1 yd = 3 ft | 1 yd$^3$ = 27 ft$^3$ |
| 1 mi = 5280 ft | 1 ft$^3$ = 7.5 gallons |
| **Weight:** | 1 cm = 0.394 in |
| 1 lb = 16 oz | 1 km = 0.621 mi |
| 1 ton = 2,000 lbs | 1 Liter = 0.264 gallons |

■ **Example 1.46** Solve the problems given the information. You need to be able to cancel out the units.

Two runners were comparing how much they had trained earlier that day. Cheryl said, "According to my pedometer, I ran 8.3 miles." Alex said, "That's a little more than what I ran. I ran 8.1 miles." How many more feet did Cheryl run than Alex?

Find the difference between the distances: 8.3 - 8.1 = 0.2 miles.

$$\frac{.2\ \cancel{mi}}{1} * \frac{5{,}280\ ft}{1\ \cancel{mi}} = \frac{1{,}056\ ft}{1} = 1{,}056\ ft$$

Cheryl ran 1,056 feet more than Alex. ■

■ **Example 1.47** There is making soup and has pint-sized containers to put it in so that she can freeze some and give it away to friends and family. She has $2\frac{3}{4}$ gallons of soup. How many pint containers will she need? (To begin, change the mixed number into an improper fraction or a decimal.)

$$\frac{11\ \cancel{gal}}{4} * \frac{4\ \cancel{qts}}{1\ \cancel{gal}} * \frac{2\ pts}{1\ \cancel{qt}} = \frac{88\ pts}{4} = 22\ pints$$

Cheree needs 22 one-pint containers for her soup.

■

In the previous examples, you should have noticed that the final units were circled in blue and in the numerators. You want to be able to cancel out your units as you proceed through the equation.

There are times when you may want to change both units. For example, your original problem may be miles per hour, but you want your final answer in feet per second.

When you cancel out your units, you want to have feet in the numerator and seconds in the denominator.

■ **Example 1.48** A professional pitcher can throw a baseball at 95 miles per hour. How fast is this in feet per second? Round your answer to the nearest feet per second.

$$\frac{95\ \cancel{\text{mi}}}{1\ \cancel{\text{hr}}} * \frac{5{,}280\ \text{ft}}{1\ \cancel{\text{mi}}} * \frac{1\ \cancel{\text{hr}}}{60\ \cancel{\text{min}}} * \frac{1\ \cancel{\text{min}}}{60\ \text{sec}} = 139.3\ \text{ft/sec}$$

A baseball pitcher can throw 139.3 ft/sec. ■

■ **Example 1.49** A 1000-foot spool of bare 12-gauge copper wire weighs 19.8 pounds. How much will 18 inches of the wire weigh in ounces?

This question asks for weight, so it should be set up as weight per length.

$$\frac{19.8\ \cancel{\text{lbs}}}{1{,}000\ \cancel{\text{ft}}} * \frac{16\ \text{oz}}{1\ \cancel{\text{lb}}} * \frac{1\ \cancel{\text{ft}}}{12\ \text{in}} = \frac{0.0264\ \text{oz}}{1\ \cancel{\text{in}}} * \frac{18\ \cancel{\text{in}}}{1} = 0.4752\ \text{oz}$$

or we could have just written one equation.

The weight of 18 inches of 12-gauge copper wire is 0.475 ounces. ■

■ **Example 1.50** You want to have a patio in your backyard. The dimensions are 16’ X 18’ and have a depth of 0.5’. How many cubic yards of concrete will be needed?

We will need to find the volume (length * width * height) of the patio before we can do any conversions. Notice that the dimensions are in feet, and the concrete comes in cubic yards.

Volume = 16’ * 18’ * 0.5’ = 144 $ft^3$

$$\frac{144\ \cancel{\text{ft}^3}}{1} * \frac{1\ \text{yd}^3}{27\ \cancel{\text{ft}^3}} = 5.3\ \text{yd}^3$$

We will need 5.3 $yd^3$ of concrete for the patio. ■

When we travel to other countries, we must exchange our money for the country currency we visit. We can do this operation at the bank or the airport. The exchange rates vary and are constantly fluctuating.


■ **Example 1.51** The currency rates as of January of 2022 were as follows:

| Currency Rates | | |
|---|---|---|
| Currency | USD per 1 Foreign | Foreign per 1 USD |
| British Pound (GBP) | 1.35806 | 0.735886 |
| Canadian Dollar (CAD) | 0.790474 | 1.265064 |
| European Euro (EUR) | 1.136193 | 0.880132 |
| Indian Rupee (INR) | 0.013463 | 74.280145 |
| Japanese Yen (JPY) | 0.008655 | 115.538761 |
| Mexican Peso (MXN) | 0.049055 | 20.38523 |
| Swiss Franc (CHF) | 1.088011 | 0.919108 |

Use this rates to find:

(a) If you have \$600, how many Japanese Yen can you get?

(b) You exchange 1,000 Indian Rupees for US dollars. How many dollars do you get?

**Solutions:**

$$(a) = \frac{\cancel{\$600}}{1} * \frac{115.538761\,\text{JPY}}{\cancel{\$1}} = 69{,}323.25\ \text{JPY}$$

$$(b) = \frac{1{,}000\ \cancel{\text{INR}}}{1} * \frac{\$0.013463}{1\ \cancel{\text{INR}}} = \$13.46$$

■

**Quick Practice - Unit Conversions:**

**Exercise 1.4.4** Answer each of the questions by using unit conversions.

(a) A bicycle moves at 15 mph. How many feet will it cover in 20 secs?

(b) Change 1.5 yards to inches. Round to the nearest inch.

(c) A faucet drips 0.25 teaspoons of water every minute. How many cups of water will be wasted in a week? Round to the nearest tenth.

(d) You have a room you want to carpet. The room has dimensions of 20 feet by 24 feet. Carpet is sold by square yards. Determine the amount of carpet that you will need to the nearest tenth. (Find the area first, which is length * width).

(e) You have 325 British Pounds (GPB). How many dollars do you have? Round to the nearest hundredth.

(f) You have $10,000 and want to convert it to Mexican Pesos (MXN). How many pesos can you get? Round to the nearest hundredth.

(g) You are traveling to Geneva, Switzerland, and the price of gasoline is 1.680 Swiss Francs per liter. How much is that in dollars per gallon? Round to the nearest hundredth.

**Solutions:**

(a) 440 feet
(b) 54 inches
(c) 52.5 cups
(d) 53.3 yards$\hat{2}$
(e) $441.37
(f) 203,852.30 MXN
(g) $6.92 per gallon

■

## 1.4.6 Exercises 1.4

**Problem 1.8** Write each of the ratios as a fraction.

(a) 84 to 36

(b) $4\frac{1}{6}$ to $3\frac{1}{3}$

(c) $1.38 to $0.69

(d) 13 inches to 2 feet

(e) Hector's total cholesterol is 249 mg/dL and HDL cholesterol is 39 mg/dL. What is his ratio of total cholesterol to HDL cholesterol? Assuming that a ratio less than 5 to

1 is considered good, what would you suggest to Hector?

**Solutions:**

(a) $\frac{7}{3}$

(b) $\frac{5}{4}$

(c) $\frac{2}{1}$

(d) $\frac{13}{24}$

(e) $\frac{83}{13}$, or as decimal number 6.4, which is high according to medical ranges. He needs to lower his total cholesterol or raise his HDL cholesterol.

**Problem 1.9** Write each rate as a fraction.

(a) 180 calories per 16 ounces

(b) 488 miles in 7 hours

(c) 8.2 pounds per 3 square inches

(d) $798 for 40 hours

**Solutions:**

(a) $\frac{45 \text{ calories}}{4 \text{ ounces}}$

(b) $\frac{488 \text{ miles}}{7 \text{ hours}}$

(c) $\frac{41 \text{ pounds}}{3 \text{ in}^2}$

(d) $\frac{\$399}{20 \text{ hour}}$

**Problem 1.10** Find the unit rate. Round to the nearest hundredth place, if necessary.

(a) A loss of 43 pounds in 16 weeks

(b) 46 heart beats in $\frac{1}{2}$ minute

(c) The bindery at a printing plant assembles 96,000 magazines in 12 hours. How many magazines are assembled in one hour?

(d) $595 for 40 hours

(e) Driving 485 miles in 6.75 hours

(f) Traveling 224 miles on 12 gallons of gasoline

**Solutions:**

(a) 2.69 pounds/week

(b) 92 beats/minute

(c) 8,000 magazines/hour

(d) \$14.88/hour

(e) 17.86 miles/hour

(f) 18.67 miles/gallon

**Problem 1.11** Compare the unit prices and determine the best deal.

(a) Brand A Chicken Noodle Soup, \$1.89 for 26 ounces, or Brand B Chicken Noodle Soup, \$0.95 for 10.75 ounces

(b) Brand C Ketchup, 40-ounce regular bottle for \$2.99 or Brand D 64-ounce squeeze bottle of Ketchup for \$4.39

(c) Brand E Breakfast cereal, 18 ounces for \$3.99 or Brand F 14 ounces for \$3.29

(d) A bag of 8 rolls for \$1.89 or a bag of 18 rolls for \$3.79

(e) Soda Pop G \$3.09 for 6 liters or Soda Pop H for \$8.29 for 16 liters.

(f) Cheese $I$ is sold by \$6.49 per lb and Cheese $J$ is sold by \$3.39 per $\frac{1}{2}$ lb.

**Solutions:**

(a) A = \$0.07/ounce, B = \$0.09/ounce; soup A is best deal

(b) C = \$0.075/ounce, D = \$0.0.69; Brand D bottle of Ketchup is the best deal

(c) E = \$0.22/ounce, F = \$0.24/ounce; cereal E is the best deal

(d) 8 rolls = \$0.24/roll; 18 rolls = \$0.21/roll, bag of 18 rolls is the best deal

(e) G = \$0.515/liter, H = \$0.518; Soda Pop G is the better deal

(f) I = \$6.49, J = \$6.78; Cheese I is a better deal

**Problem 1.12** Solve the applications using proportions.

(a) Dylan and David are planning a backpacking trip in Yosemite planning a trip to Yosemite National Park. On their map, the legend indicates that 1.2 centimeters represent 2 miles. How long is their trip if the route measures 10.6 centimeters on the map? Round your answer to the nearest tenth of a mile.

(b) US Highway 199 had a landslide where as much as 3 000 cubic yards of material fell on the road, reportedly requiring about 200 large dump trucks to remove. A week earlier, 40 000 cubic yards of material fell on Highway 96. Estimate the number of dump trucks needed for that slide rounded to the nearest whole number. *Associated Press-Times-Standard 03/09/10 Another highway closed by slide*

(c) In Australia, penalties on ships causing oil spills are approximate 1.75 million, equivalent to about 1.64 million US dollars. After an oil tanker was grounded on a coral reef, Australian officials considered raising the fine to 10 million Australian dollars. What will the new fine be in US dollars? Round your answer to the nearest hundredth of a million dollars. *Associated Press-Times-Standard 04/13/10 Ship that leaked oil on Great Barrier Reef removed.*

(d) Jesse's car gets 30 miles per gallon of gas. If Las Vegas is 285 miles away, how many gallons of gas are needed to get there and then home? If gas is \$3.09 per gallon, what is the total cost of the gas for the trip?

(e) At the laundromat, Lucy changed \$12.00 into quarters. How many quarters did she get?

(f) At the gym, Carol takes her pulse for 10 sec and counts 19 beats. How many beats per minute is this? Has Carol met her target heart rate of 140 beats per minute?

**Solutions:**

(a) 17.7 miles

(b) 2,667 loads

(c) $9.37 million US dollars

(d) 19 gallons; $58.71

(e) 48 quarters

(f) 114 beats/minute; Carol has not met her target heart rate.

**Problem 1.13** Solve using unit conversions.

(a) How many gallons is 32 fluid ounces?

(b) Ramona is planning a trip to Orlando, which is 485 miles away. Her car gets 32 miles per gallon and the cost of gas is $2.87. How much will it cost her for gas to travel to Orlando?

(c) An airplane flies at 565 miles per hour. How many yards per minute does a plane fly? Round to the nearest tenth.

**Solutions:**

(a) 0.25 gallons

(b) $43.50

(c) 16,573.3 yards/min

## 1.5 Geometry

Everywhere you look, there are shapes. Buildings, roads, vehicles, road signs, and packages are just a few items we could describe using a geometric shape. Sometimes we could put more than one shape together to form another unique shape. In this section, we will take a look at some of these shapes, but more importantly, we will do some calculations with their formulas.

### 1.5.1 Length, Perimeter, Area, and Volume

**Definition 1.5.1 length** measures the distance in a line from a starting point to an ending point. It can be measured in inches, feet, miles, centimeters, kilometers, etc.

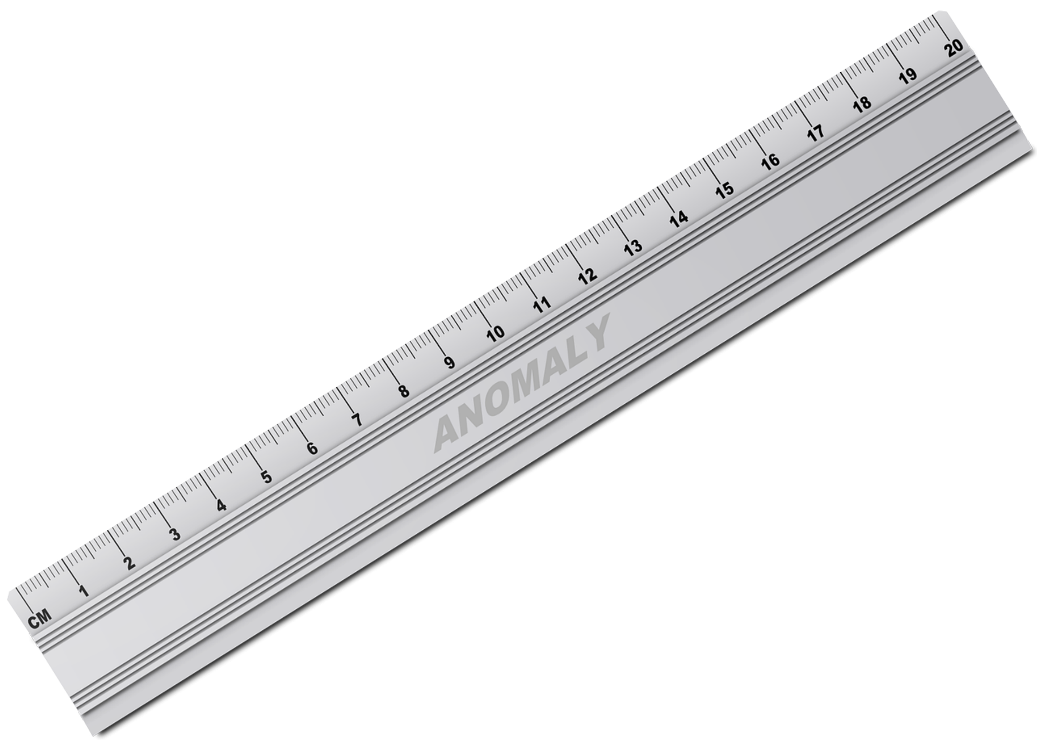
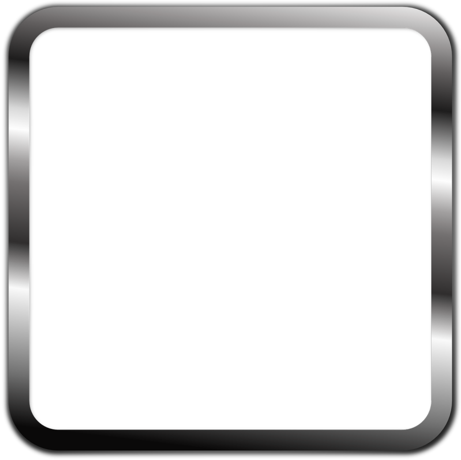



Figure 1.6: We use elements like rulers to measure length.

We can find the measure around the boundary of flat objects, called the perimeter, using these measurements. When we talk about a circle, we call this the circumference.

**Definition 1.5.2 perimeter** is the measurement that forms a boundary around a shape. We calculate the measure by adding the length of the sides. Examples of this can include a frame around a picture or a fence around a yard. We can measure it in inches, feet, miles, centimeters, kilometers, etc.

**Example 1.52** Given the frame where each side is 12 cm, find the perimeter.



Since we know the length of each of the sides is 12 cm, we can calculate the perimeter P = 12 cm + 12 cm + 12 cm + 12 cm = 48 cm. Since this is a square, another way that we could have calculated this would be, P = 4 * 12 = 48 cm. ∎

How would we know how much paint to buy if we wanted to paint our living room? We would not want to guess because we don't want to have to buy more paint and we can't return the extra paint. Therefore, we can find the living room area and determine the number of gallons of paint needed to complete the job.

**Definition 1.5.3** **Area** measures the amount of flat space a shape covers, for example, the area of a floor or the size of a city. We measure the area in square units such as square inches (inches$^2$), square centimeters (cm$^2$), square feet (feet$^2$), etc.

**Example 1.53** How much area can we cover in the pavement if we have 250 square tiles that are 8 cm in length?

If the length of the square tile is 8 cm, we can calculate the area by multiplying the length of the sides. Since the area per tile is 8 cm * 8 cm = 64 cm$^2$ and since we have 250 tiles, we can cover 250 * 64 cm$^2$ = 16,000 cm$^2$ of space. ∎

**Definition 1.5.4 volume** measures the amount of three-dimensional space that a substance or an object occupies within that space.

Examples of volumes include the amount of water in the bathtub, the number of gallons of gas in your gas tank, or the size of brick used to build a house. We measure volumes in cubic units (units$^3$), like cubic inches (in$^3$), cubic centimeters (cm$^3$), cubic feet (ft$^3$), etc.

Figure 1.7: Solids are special geometric figures with a volume that can be easily calculated

#### Quick Practice - Measurements

**Exercise 1.5.1** Answer if these would be linear measure, perimeter measure, square measure, or cubic measure.

(a) Kellogg's Frosted Flakes Box
(b) Christmas lights around the house
(c) Length of a curtain rod
(d) Can of Coke
(e) Rain gutters around a building
(f) Carpeting your bedroom
(g) Roof size of your doghouse
(h) Height of a tree

**Solutions:**

(a) Cubic Units
(b) Perimeter Units
(c) Linear Units
(d) Cubic Units
(e) Perimeter Units
(f) Area Units
(g) Area Units
(h) Linear Units

### 1.5.2 Basic Geometric Shapes and Their Formulas

Below are the figures showing the basic geometric shapes and their formulas. We will be using these formulas to solve problems involving perimeter and area.

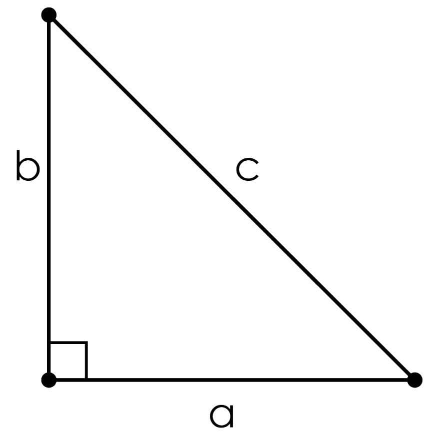



Figure 1.8: Right Triangle with legs a, b and hypotenuse c

A **triangle** is a figure created by intersecting three lines, all the internal angles add to 180°. One special case is the right triangle. This is when one of the angles measures 90°. The right angle is recognized by the "box" and the hypotenuse is always opposite the right angle.

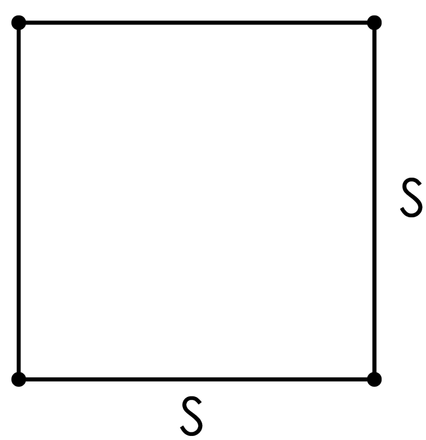



Figure 1.9: Square with sides "s"

A **square** is a figure created by intersecting two sets of parallel lines where each side has the exact measurement, and each angle is 90°.


Figure 1.10: Rectangle with length "L" and width "W"

A **rectangle** is a figure created by intersecting 2 sets of parallel lines where the opposite sides have the same measurement and each angle is 90°.

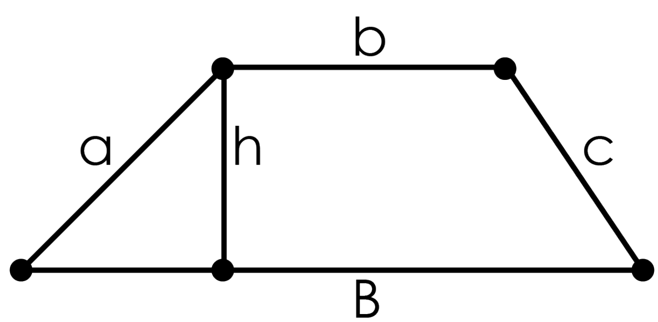



Figure 1.11: Trapezoid with bases "b" and "B", legs "a" and "c", and height "h"

A **trapezoid** is a four-sided figure, a quadrilateral, that has two parallel sides that are called bases and two non-parallel sides that are called legs. We name the smaller base's length as b and the length of the larger base as B. The altitude (or the height) is the distance between the two bases. The total degree measurement is 360°. There are three types of trapezoids: Isosceles, Scalene and Right Trapezoids.

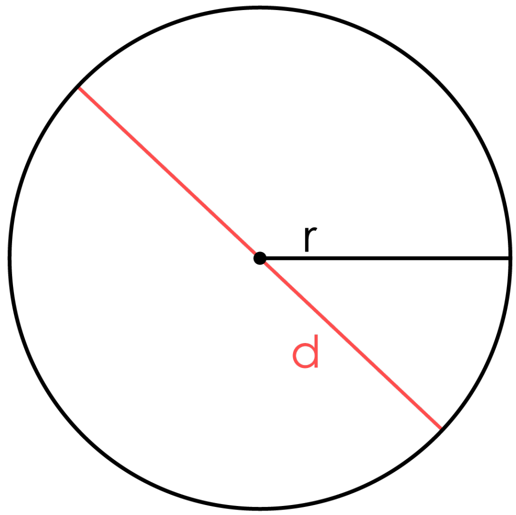



Figure 1.12: Circle with radius "r" and diameter "d"

A **circle** is a shape consisting of all points in a plane that are at the same distance from a given point called the center. The radius is $\frac{1}{2}$ the measurement of the diameter, and the total number of degrees in a circle is 360°.

The table below gives the formulas to calculate the perimeter and area for the geometric shapes described above.

| **Geometric Shape** | **Perimeter (Units)** | **Area (Units$^2$)** |
|---|---|---|
| Right Triangle | P = a + b + c | $A = \frac{1}{2}b * h$ |
| Square | P = 4 * s | $A = s^2$ |
| Rectangle | P = 2L + 2W | A = L * W |
| Trapezoid | P = a + b + c + B | $A = \frac{b+B}{2} * h$ |
| Circle | Circumference = $2\pi$r or $\pi$d | $A = \pi r^2$ |

Notice that we can calculate the perimeter and area for each geometric shape by using their measurements.

■ **Example 1.54** The length of a rectangle formed in the patio of a building is 32 meters, and the width is 20 meters.

Find the perimeter and the area that the patio covers.

**Solution:**

First, determine the dimension to use from the problem. The calculations require the length $L$ and width $W$ for both perimeter and area. Let's calculate the perimeter, $P$.

$$
\begin{aligned}
P &= 2L + 2W \\
&= 2(32) + 2(20) \\
&= 64 + 40 = 104
\end{aligned}
$$

Now let's calculate the area, $A$.

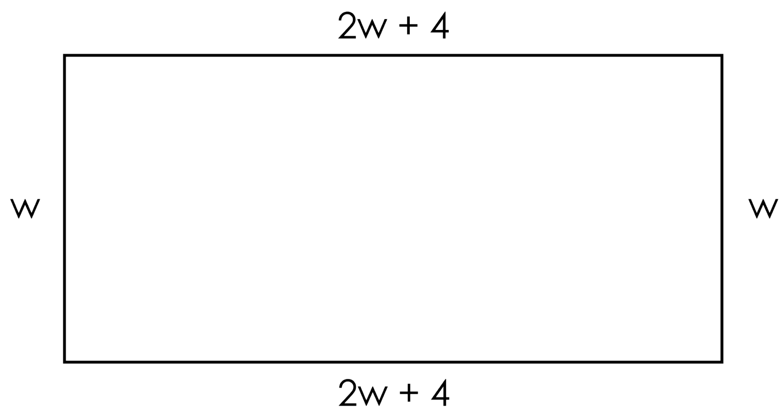
$$
\begin{aligned}
A &= L * W = 32 * 20 \\
&= 640
\end{aligned}
$$

The perimeter is 104 m, and the area is 640 m$^2$ that the patio covers.

■

■ **Example 1.55** The length of a rectangle is four centimeters, more than twice the width. The perimeter is 32 centimeters. Find the height and width.

Let w represent the width and let 2w + 4 represent the length since it is four centimeters more than twice the width.




**Solution:**

First, set up the problem using the formula:

$$\text{Perimeter} = 2L + 2W$$

Now we substitute the numbers for the length and width.

$$\begin{aligned} 32 &= 2(2W+4)+2(W) \\ &= 4W+8+2W && \text{Multiply the 2 inside the parenthesis} \\ &= 6W+8 && \text{Add the W's together} \\ 24 &= 6W && \text{Subtract the 8 on both sides} \\ 4 &= W && \text{Divide by 6 on both sides} \end{aligned}$$

To solve for the length, substitute 4 in for W and solve for L.

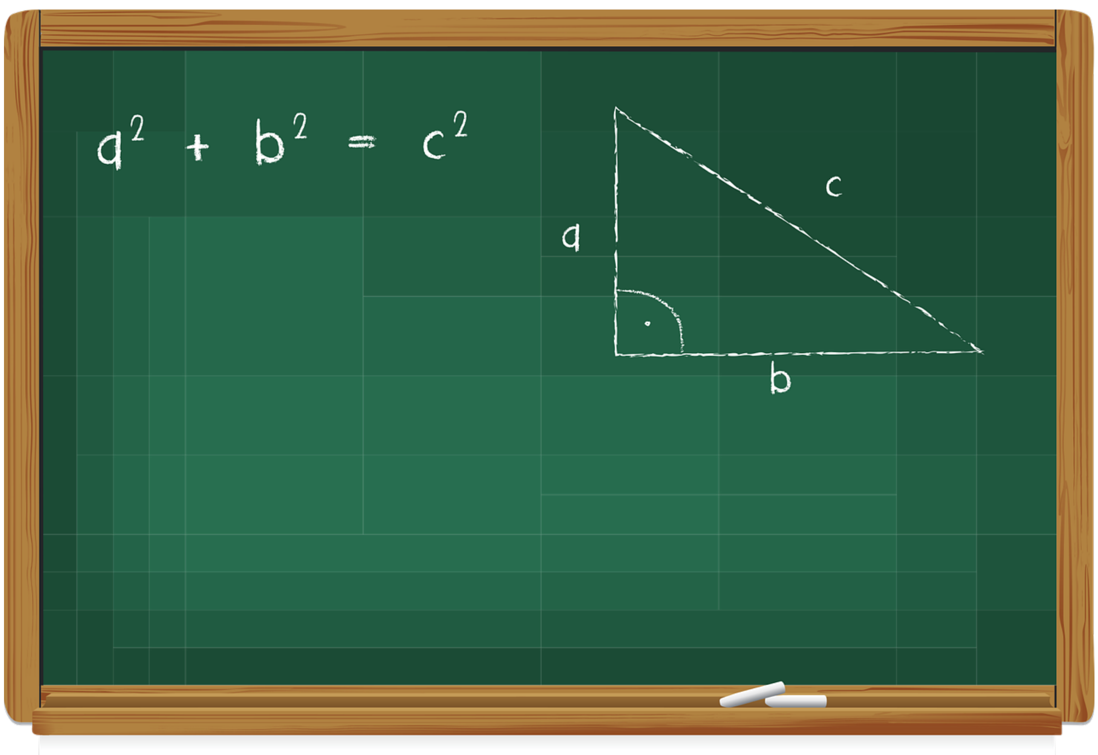
$$\begin{aligned} 32 &= 2L+2(4) && \text{Use the perimeter formula with } P=32 \text{ and } W=4 \\ &= 2L+8 && \text{Multiply } 2*4 \\ 24 &= 2L && \text{Subtract 8 from both sides} \\ 12 &= L && \text{Divide 2 from both sides} \end{aligned}$$

The width is 4 centimeters, and the length is 12 centimeters. ■

Triangles have three sides and a measure of 180 °.

- **Equilateral:** "equal"-lateral (lateral means side), so all sides are equal.
- **Isosceles:** means "equal legs", so two legs are equal (because we have two legs, right?).
- **Scalene:** means "uneven" or "odd", so no sides are equal.

There are several types of triangles, but we will concentrate on the right triangles. That means that they have a right angle, which is indicated by a box in one of the corners; this means the angle with the box measures 90 °.

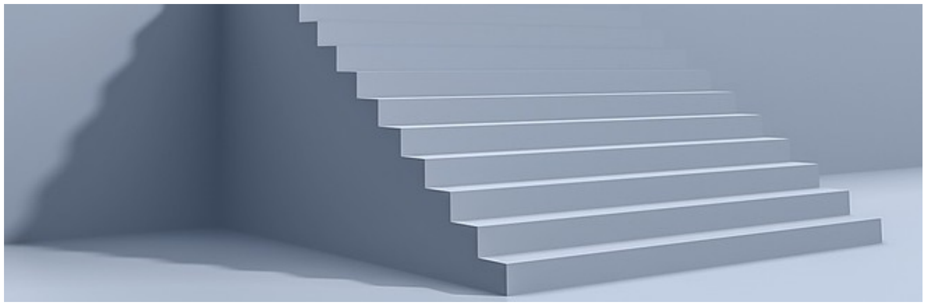



To determine the measurements of the legs or hypotenuse of a right triangle, we can use the Pythagorean Theorem, which states that the sum of the squares of the legs is equal to the square of the hypotenuse.

$$a^2 + b^2 = c^2.$$

The hypotenuse is always the largest leg, and it is opposite from the right angle. The Pythagorean Theorem **only** works for a right triangle.

■ **Example 1.56** Solve the problem using the Pythagorean Theorem.

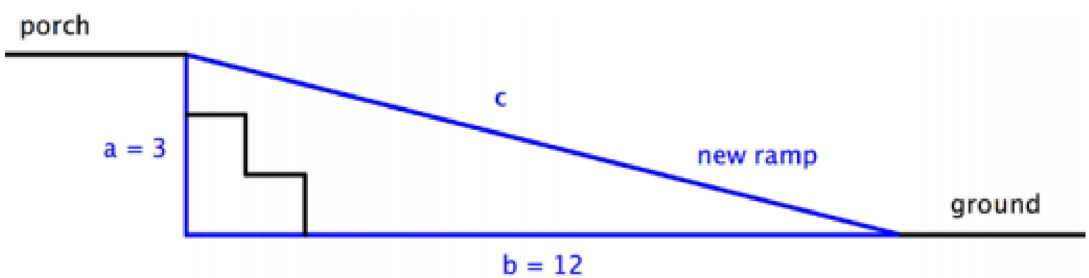


The house owners want to convert a stairway leading from the ground to their back porch into a ramp. The porch is 3 feet off the ground, and due to building regulations, the ramp must start 12 feet away from the base of the porch.

How long will the ramp be? Use a calculator to find the square root and round the answer to the nearest tenth.

**Solution:**

Let a = 3, b = 12, c = c




Set up the problem using the Pythagorean Theorem: $a^2 + b^2 = c^2$

| | |
|---|---|
| $3^2 + 12^2 = c^2$ | Substitute in the numbers |
| $9 + 144 = c^2$ | Square the 3 and 12 |
| $153 = c^2$ | Add the numbers together |
| $12.4 = c$ | Take the square root of both sides |

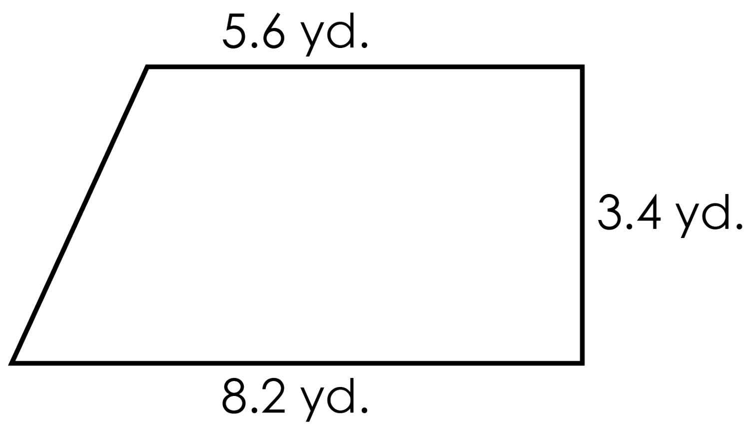
The ramp will be 12.4 feet long.

■

■ **Example 1.57** Solve the problem using the area formula for a trapezoid.

Vinny has a garden shaped like a trapezoid. The trapezoid has a height of 3.4 yards, and the bases are 8.2 and 5.6 yards. How many square yards will be available to plant?

Let b = 5.6, B = 8.2, h = 3.4




Set up the problem:

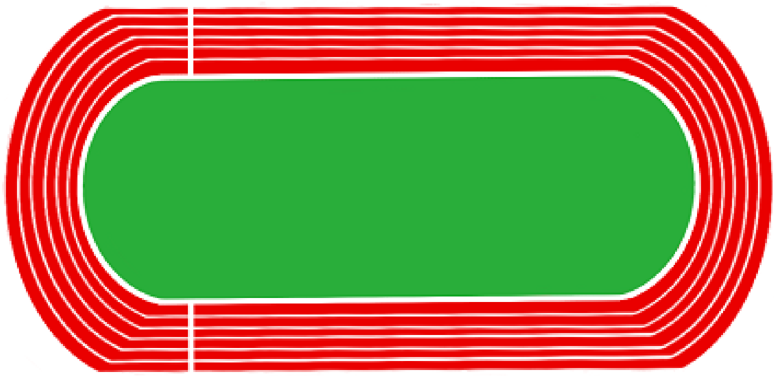
$$A = \frac{b+B}{2} * h$$

$$\begin{aligned} A &= \frac{5.6+8.2}{2} * 3.4 && \text{Substitute in the numbers} \\ &= \frac{13.8}{2} * 3.4 = 23.46 && \text{Add the numerator numbers and simplify} \end{aligned}$$

Vinny has 23.46 yards$^2$ in which he can plant. ■

Sometimes we can combine shapes, and we need to pull them apart to calculate the areas for each one separately.

■ **Example 1.58** A high school track is shaped like a rectangle with a semicircle (half a circle) on each end. The rectangle has a length of 105 meters and a width of 68 meters.

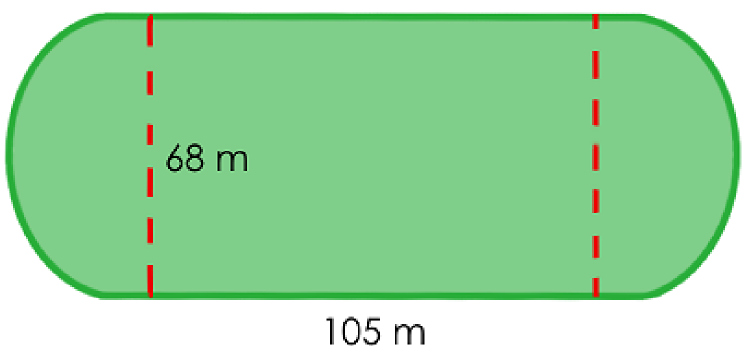


Find the area enclosed by the track. Round your answer to the nearest hundredth.

**Solution:**

We will break the figure into a rectangle and two semicircles. The area of the figure will be the sum of the areas of the rectangle and the two semicircles (which add to a full circle). Therefore, the total area is the sum of the two figures.

$$\text{Total Area} = \text{Rectangle} + \text{Circle} = L * W + \pi r^2$$




$$\text{Total Area} = \text{Rectangle} + \text{Circle} = 105 * 68 + 3.14 * (34)^2 \text{ (the radius is } \frac{1}{2} \text{ of 68 m)}$$

$$\text{Total Area} = 7{,}140 + 3{,}629.84 = 10{,}769.84$$

The area enclosed by the track is 10,769.84 meters$^2$.

■

**Quick Practice - Area and Perimeter**

**Exercise 1.5.2** Solve the problems given the following information. Draw a picture.

(a) John puts the base of a 13-ft ladder 5 feet from the wall of his house. How far up the wall does the ladder reach?

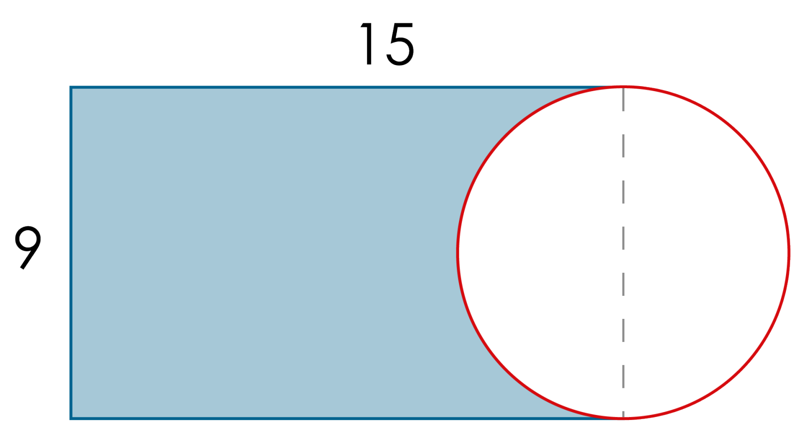
(b) The width of a rectangle is six less than twice the length. The perimeter is 18 centimeters. Find the length and width.

(c) The area of a triangular church window is 90 meters $^2$. The base of the window is 15 meters. What is the window's height?

(d) Find the area of a trapezoid whose height is 5 feet and whose bases are 10.3 and 13.7 feet.

(e) A circular sandbox has a radius of 2.5 feet. Find the circumference and the area.

(f) Find the area to the nearest hundredth for the shaded region in the figure below.




**Solutions:**

(a) 12 feet
(b) L = 5cm, W = 4cm
(c) 12 meters
(d) 60 feet$^2$
(e) C = 15.7 feet, A = 19.625 feet $^2$
(f) 166.79 units$^2$

### 1.5.3 Geometric Solids and Their Formulas

When we are looking at geometric solids, we are looking at three-dimensional figures where they have a length, width, and height. The flat side of a geometric surface is called the face.

Here we can see some examples of solids and real-world examples.

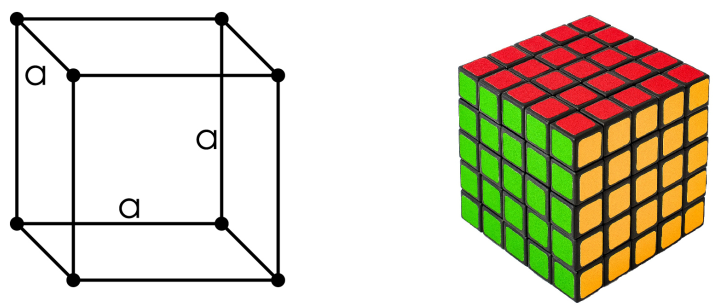



Figure 1.13: Cube with sides "a" and a Rubik's Cube

The cube is a shape you see every day. For example, sugar cubes, ice cubes, dice, and a Rubik's cube are just a few items with a cube shape.

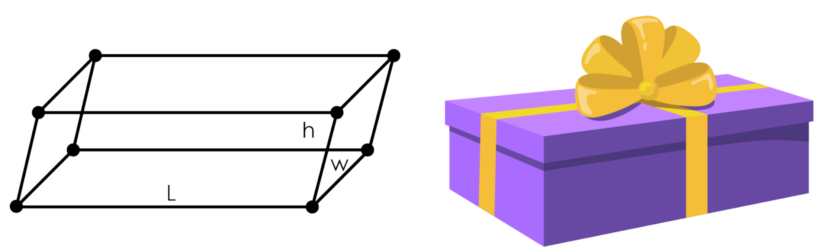



Figure 1.14: Rectangular Prism with width "w", length "L", height "h" and a shoe box

Rectangular prisms are a standard shape you see every day. For example, your phone and laptop are rectangular prisms, and so are shoe boxes, tissue boxes, and a hard copy of your math book.

Pyramids in Egypt are great examples of this solid in the real world.

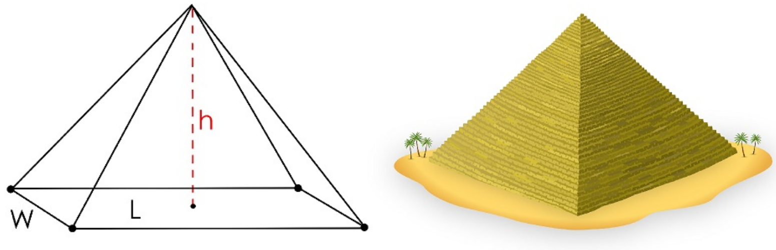



Figure 1.15: Pyramid with width "w", length "L", height "h" and The Great Pyramid

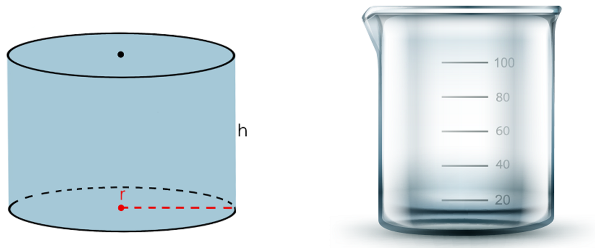



Figure 1.16: Cylinder with height "h", radius "r" and a graduated cylinder for labs.

We use cylindrical shapes every day as they come in vases, jars, soda cans, and even candles.

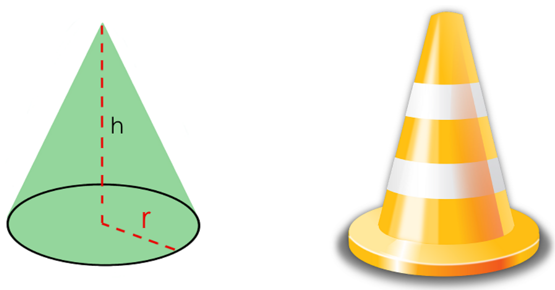



Figure 1.17: Cone with height "h", radius "r" and a traffic cone

Cone-shaped items can be seen when you eat an ice cream cone when you are stopped with the traffic cones blocking a lane of traffic, or you may have a home built with a turret design.


Figure 1.18: Australia's daily greenhouse gas emissions

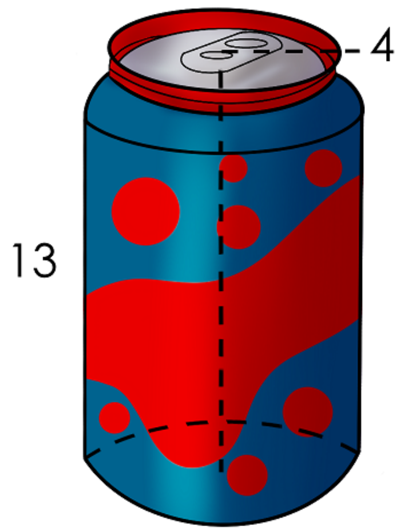
The sphere is recognizable by basketballs and baseballs, but you also see that apples, oranges, and light bulbs have the same shape.

When you look at a shoe box, you will notice that it is rectangular and three-dimensional. Looking at the formula sheet below, the cube, rectangular pyramid, and pyramid have a base of a rectangle. The cylinder and a cone each have as a bottom a circle. The sphere is unique as it is circular, but the points on the sphere are the same distance from the middle of the sphere. That is what makes it round like a ball or a globe.

| **SOLID NAME** | **SURFACE AREA (Units)** | **VOLUME (Units$^3$)** |
|---|---|---|
| Cube | SA = $6a^2$ | V = $a^3$ |
| Rectangular Prism | SA = 2(Lw + Lh + wh) | V = Lwh |
| Pyramid | | V = $\frac{Ah}{3}$ |
| Cylinder | SA = $2\pi rh + 2\pi r^2$ | V = $\pi r^2 h$ |
| Cone | | V = $\frac{1}{3}\pi r^2 h$ |
| Sphere | SA = $4\pi r^2$ | v = $\frac{4}{3}\pi r^3$ |

We will be using these formulas to solve problems.

■ **Example 1.59** A soda can has a base radius of 4 centimeters, and the height is 13 centimeters. Find the volume of the can.




**Solution:**

First, we set up the problem. In this case, we will imagine we can approximate the can of soda via a cylinder. Hence

$$\text{Volume of a cylinder} = \pi r^2 h$$

Then we substitute in the dimensions.

$$\text{Volume} = 3.14 * (4)^2 * 13$$

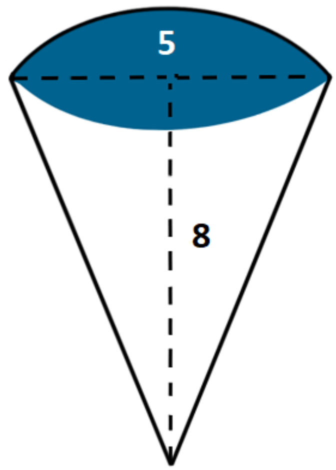
Calculate the square of the radius and multiply.

$$\text{Volume} = 3.14 * 16 * 13 = 653.12 \text{ cm}^3$$

The volume of the can of soda is 653.12 cubic centimeters. ■

We need to be very careful when writing units of area or volume in terms of units of length, as for sure, we will need powers of the particular unit. Other units like the liter or gallons can also be considered units of volume and have a particular translation to the cubic units.

■ **Example 1.60** Katrina's favorite pub serves french fries in a paper wrap shaped like a cone.




What is the volume of a conic wrap that is 8 inches tall and 5 inches in diameter? Round the answer to the nearest hundredth.

**Solution:**

We use the formula for the volume of the cone:

$$\text{Volume} = \frac{1}{3}\pi r^2 h,$$

We substitute our quantities in the radius and height variables.

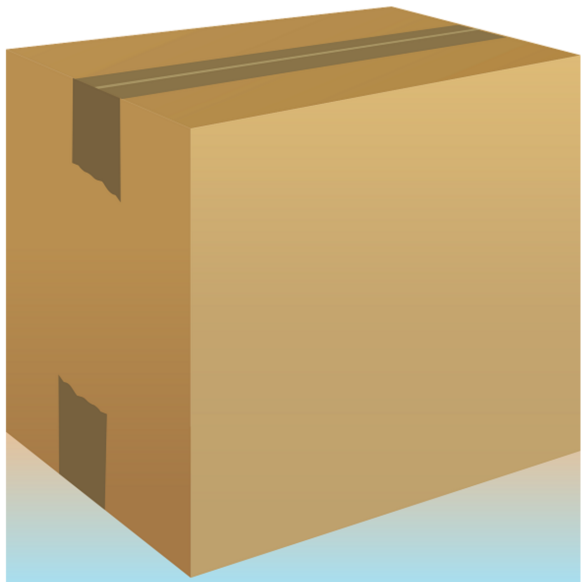
$$\text{Volume} = \frac{1}{3} * 3.14 * (2.5)^2 * 8$$

After computing the square, we multiply and obtain the volume.

$$\text{Volume} = \frac{1}{3} * 3.14 * 6.25 * 8 = 52.33 \text{ inches}^3$$

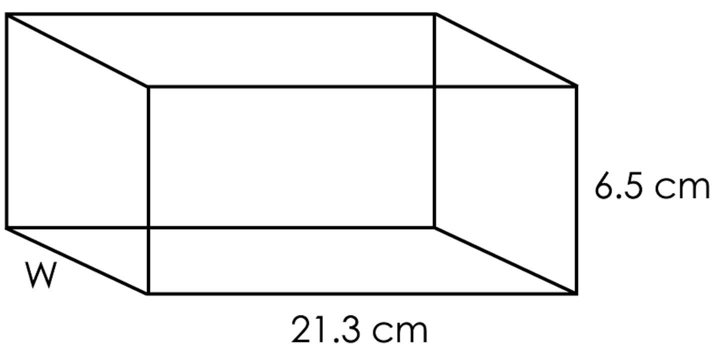
The volume of the paper cone is 52.33 inches$^3$. ■

■ **Example 1.61** A carton, or a box, has a rectangular shape and has a volume of 3,350.49 $cm^3$. If the length is 21.3 cm and the height is 6.5 cm, what is the width of the carton to the nearest tenth?



**Solution:**

We can see the box as a solid and use the dimension to solve the problem.




Set up the problem by using the volume formula.

$$\text{Volume} = l * w * h$$

We replace the volume and the variables for length and height but leave the "w" alone since that is what we need to find.

$$3{,}350.49 = 21.3 * w * 6.5.$$

Simplify the problem

$$3{,}350.49 = 138.45 * w$$

Divide by 138.45 on both sides to obtain $24.2 = w$.

The width of the carton is 24.2 cm. ■

We know that volume is the amount of something that takes up space in an object. We also know how to use proportions and unit conversions. Let's combine these two topics to see how we may use them in our lives.

■ **Example 1.62** Gracie is going to carpet her living room. The dimensions are 15' x 18'. The carpet is priced at $7.67 per square yard. Determine the amount of area that needs to be carpeted and the total cost of the carpet.

**Solution:**

Find the area of the room: 15 ft * 18 ft = 270 $ft^2$.

Set up the conversion: $\dfrac{270\ \text{ft}^2}{1} * \dfrac{1\ \text{yd}^2}{9\ \text{ft}^2} = 30\ \text{ft}^2$

Find the price: 30 $ft^2$ * $7.67 = $230.10

The total area that Grace will carpet in her living room is 30 $ft^2$, and it will cost her $230.10.

■

■ **Example 1.63** A website says that you'll need 48 fifty-pound bags of sand to fill a sandbox that measures 8 ft x 8 ft x 1 ft. How many bags would you need for a 6 ft x 4 ft x 1 ft sandbox?

**Solution:**

Volume of large sandbox: 8 * 8 * 1 = 64 $\text{ft}^3$

Volume of small sandbox: 6 * 4 * 1 = 24 $\text{ft}^3$

We can use a proportion to solve this problem.

$$\frac{\text{48 bags}}{\text{64 ft}^3} = \frac{\text{B}}{\text{24 ft}^3}$$

$$64B = 1,152 \Longleftrightarrow B = 18$$

The smaller sandbox will need 18 bags of sand.

■

Flat figures can be measured in terms of area, but also we can establish the length of their sides via the perimeter. Solids work similarly; we can measure their volume and the area created by the surface delimiting the objects.

### 1.5.4 Surface Area of Geometric Solids

**Definition 1.5.5 Surface area** describes the area of a three-dimensional figure.

If you want to paint the outside of your house, you need to know the surface area to purchase enough paint. A company that manufactures boxes would need to see the surface area so that they could create the box with very little waste. The units for the surface area are the same as with area; the units would be units$^2$.

■ **Example 1.64** Meko knows that the surface area of her hot air balloon is 16,277.76 ft$^2$, but she wants to see the length of the radius.

Set up the problem using the surface formula.

$$\text{Surface area} = 4\pi r^2$$

Now we set the known value of the surface area to be equal to the formula:

$$16{,}277.76 = 4 * 3.14 * r^2$$

After substituting the numbers we multiply and simplify values:

$$16{,}277.76 = 12.56 * r^2$$

Finally, we solve for $r^2$ to get $r$

$$1296 = r^2 \quad \rightarrow 36 = r$$

The length of the radius of Meko's hot air balloon is 36 feet. ■

**■ Example 1.65**

Tyrone is wrapping a rectangular gift box that has the dimensions of a length of 26 inches, a width of 16 inches, and a height of 4 inches. Find the surface area.


Again, start by setting up the problem with the surface area.

$$\text{Surface area} = 2(lw + lh + wh)$$

Substitute the numbers into the corresponding dimensions

$$\text{Surface area} = 2(26 * 16 + 26 * 4 + 16 * 4)$$

Multiply, add, then simplify

$$\text{Surface area} = 2(416 + 104 + 64) = 2(584)$$

$$\text{Surface area} = 1{,}168$$

The surface area of Tyrone's gift box is 1,168 inches$^2$. ■

#### Quick Practice - Surface Area

Solve the problems given the following information. Draw a picture.

**Exercise 1.5.3**

(a) Zel was getting ready for a party and bought a chunk of cheese in a rectangular pyramid. The volume is 18.67 $in^3$. If the length is 4 inches and the height is 7 inches, what is the width?

(b) A rectangular moving van has a length of 16 feet, a width of 8 feet, and a height of 8 feet. Find the volume and the surface area.

(c) A barber shop pole has a cylinder shape. The surface area is 508.68 $in^2$. The diameter is 6 inches. Find the height.

(d) The statue's base is a cube with sides 2.8 meters long. Find the volume and the surface area to the nearest hundredth.

**Exercise 1.5.4** A golf ball has a radius of 4.5 centimeters. Find the volume.

**Exercise 1.5.5** Find the area of the shaded region in 1.19. Round to the nearest tenth. ■

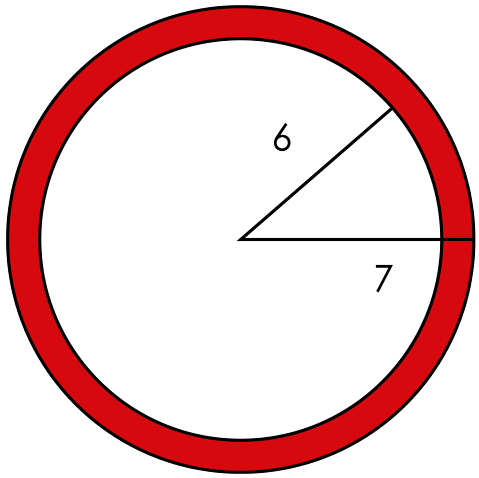



Figure 1.19:

**Exercise 1.5.6** A capsule is shown below in figure 1.20. If the radius of the spherical ends is 6 inches, find the volume of the solid below. Round to the nearest hundredth. ■

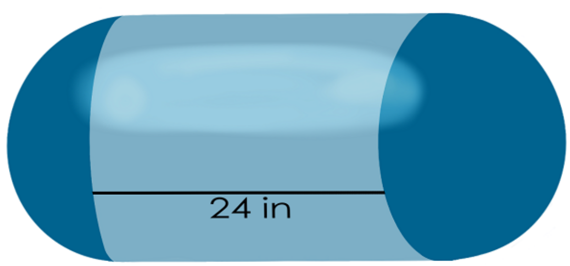



Figure 1.20:

**Solutions:**

(a) 2 inches
(b) V = 1,024 $ft^3$, SA = 640 $ft^2$
(c) 24 inches
(d) V = 21.95 $m^3$, SA = 47.04 $m^2$

(1.5.4) 381.54 $cm^3$
(1.5.5) 40.82 $units^2$
(1.5.6) 3,617.28 $in^3$

### 1.5.5 Exercises 1.5

For each of the geometric problems, draw a picture and solve it.

**Problem 1.14**

(a) Pat wants to paint one wall of his attic. The wall is shaped like a trapezoid with a height of 8 feet and bases of 20 feet and 12 feet. The cost of painting one square foot of wall is about \$0.05. Approximately how much will it cost for Pat to paint the attic wall?

(b) Paula wants to fence around her triangular flowerbed. The sides of the flowerbed are 6 feet, 8 feet, and 10 feet. The fence costs \$10 per foot. How much will it cost for Paula to fence her flowerbed?

(c) Use unit conversions to solve. Tomiko is putting a rectangular pool in her backyard. The dimensions are 10' x 20' x 5.5'. It will cost \$6.00 per 1,000 gallons to fill the pool. How much will it cost to fill the pool with water?

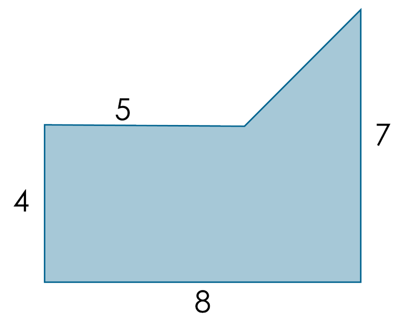
(d) What is the volume of a cone-shaped silo 50 feet tall and 70 feet across at the base? Round to the nearest hundredth.

**Problem 1.15** A sailboat has a giant sail in the shape of a right triangle. The longest edge of the sail measures 17 yards, and the bottom edge of the sail is 8 yards. How tall is the sail?

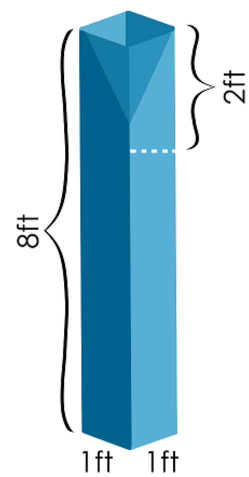
**Problem 1.16** A box of tissues with sides 4.5, 7.5, and 10 inches long, respectively. Round to the nearest tenth. Find the Volume.

**Problem 1.17** A circular saw has a circumference of 37.68 inches. What is the radius?

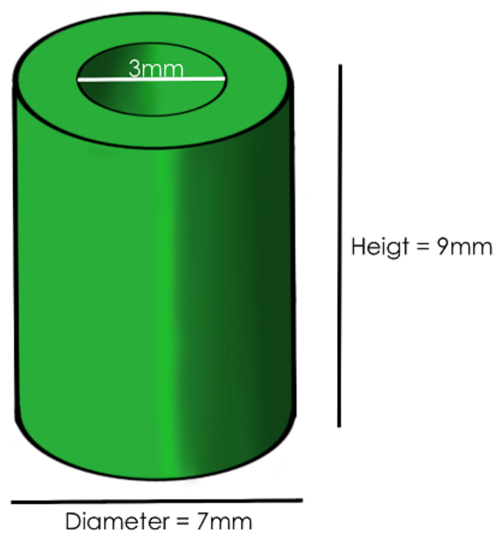
**Problem 1.18** Find the area of the irregular shape. Round to the nearest tenth.




**Problem 1.19** A sculptor carves a rectangular prism out of a solid piece of wood. Then, at the top, she hollows out an inverted pyramid. The solid, and its dimensions, are shown below. What is the volume of the finished piece, rounded to the nearest tenth?



**Problem 1.20** A machine takes a solid cylinder with a height of 9 mm and a diameter of 7 mm and bores a hole through it. Find the volume of the solid to the nearest tenth if the hole it creates has a diameter of 3 mm.




**Solutions:**

(1.14-a) \$6.40

(1.14-b) \$240.00

(1.14-c) \$49.50

(1.14-d) 64,108.33 $ft^3$

(1.15) 15 yards

(1.16) 337.5 $in^3$

[1.17) 6 inches

(1.18) 36.5 $units^2$

(1.19) 7.3 $ft^3$

(1.20) 282.6 $mm^3$

## 1.6 Projects

### 1.6.1 Backyard Renovation

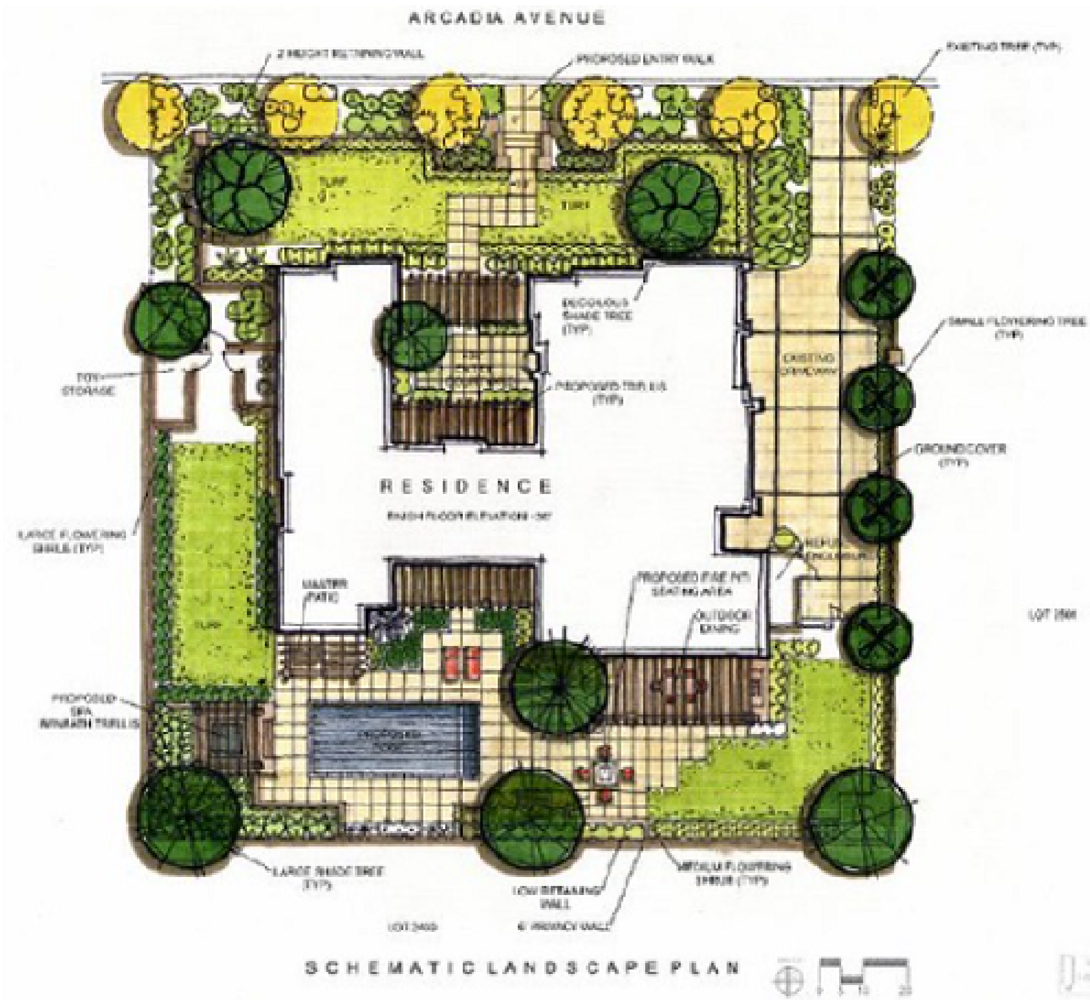



Figure 1.21: Photo Credit: Landscape Design Advisor

This project will combine several of the topics covered in this chapter.

You have just bought your dream home, and you want to do some landscaping and decorating in the backyard. Your wish list for the space includes:

1. Sod the entire area
2. Install a hot tub on a cement slab
3. Plant a small vegetable garden

4. Paint a new table

5. Plant 2 trees fruit trees, two shade trees

6. Buy 16 vegetable plants

7. fence the backyard (not the back of the house)

Here are the dimensions that you will need to use:

1. The backyard is in a rectangular shape and has the dimensions of 40’ x 60’.

2. The hot tub dimensions are 84" x "64" x 29".

3. The slab for the hot tub will be 112" x 132" x 4" thick.

4. The garden is in the shape of a trapezoid with bases of 12.5’ and 18.3’ and a side of 15’.

5. You have a circular table that sits on a cylindrical post. The table has a diameter of 30". The cylinder has a radius of 8" and a height of 30".

6. The back of the house measures 30’ and is spaced evenly between the sides of the yard.

Figure 1.22: Photo Credit: cseeman

Here are the costs that you will need to use:

1. Grass is \$4.99 per $ft^2$
2. The hot tub is regularly \$7,770, but there is a sale for 20% off
3. The cost to fill the hot tub will be \$9.00 per 1,000 gallons
4. The cement to pour the slab will cost \$4.99 per $ft^2$
5. The garden soil is \$7.97 a bag for 1.5 $ft^3$
6. A quart of paint will cover 100 $ft^2$, and it will cost \$12.98
7. The fence will cost \$187.62 for a 30" x 30" section, one section will be for a gate, and that width will be 28" in length and will cost \$29.98
8. You need to purchase two fruit trees and two shade trees to plant in your backyard
9. You need to purchase 16 vegetable plants (not all one kind)
10. The tax on these items is 7.125%

**Project Directions:**

1. You are to design your backyard using your wish list. We must include all of these items, and they can be placed wherever you want them. It needs to be drawn on graph paper so that you can make a scale of your drawing.

2. To determine the cement, garden soil, etc., calculations need to be made before you can determine the number of materials to purchase. This work will need to be turned in with your project.

3. You are to keep track of all your expenses on a sheet of paper or an excel spreadsheet. Must group each of the items. For example, put all of the hot tub items together, then all the trees together, etc. Don't forget the tax!

4. When you find your fruit and shade trees, take a snip of the picture and put them in your report, along with the price for each one. You need to do the same for the vegetable plants.

5. You can add anything else to your wish list.

You will be turning in a report, no longer than two pages, double spaced and written in 12 font, that will include answers to the questions:

1. How did you approach this project? Did you start with the big items first or the garden, etc.?

2. Are you surprised at the amount of math needed before you made any purchases?

3. What would you change if you had to redo this project?

4. Do you see how this can be useful when you own a home and want to do your landscaping or renovations?

The report needs to include the additional pages:

1. Your backyard design (on graph paper)

2. Spreadsheet with your budget

3. Pictures of your trees and vegetable plants with the prices

4. All calculations (do not type these)

### 1.6.2 Covid and Percentages

**Instructions:**

In this project, we take a look at the numbers of Covid 19 infections and deaths, drawing a comparison between different counties in the state of Georgia.

Our source will be the websites:

- `https://dph.georgia.gov/covid-19-daily-status-report`
- `https://covid.cdc.gov/covid-data-tracker/#county-view`

We can contrast information between information given by the CDC and the State of Georgia. To compare the growth or decay of the infection rates, we will use the concepts of **Absolute Change** and **Relative Change**.

First, let's take a look at the Georgia site and the visual information we can obtain. We can examine the information visually using the Georgia map on this website. We also get data in the Counties in categories as positive cases, cases by 100k, deaths, and hospitalizations. Considering numbers in the race, sex and others. You will use this data and the questions below to write a **report** or **paper**.

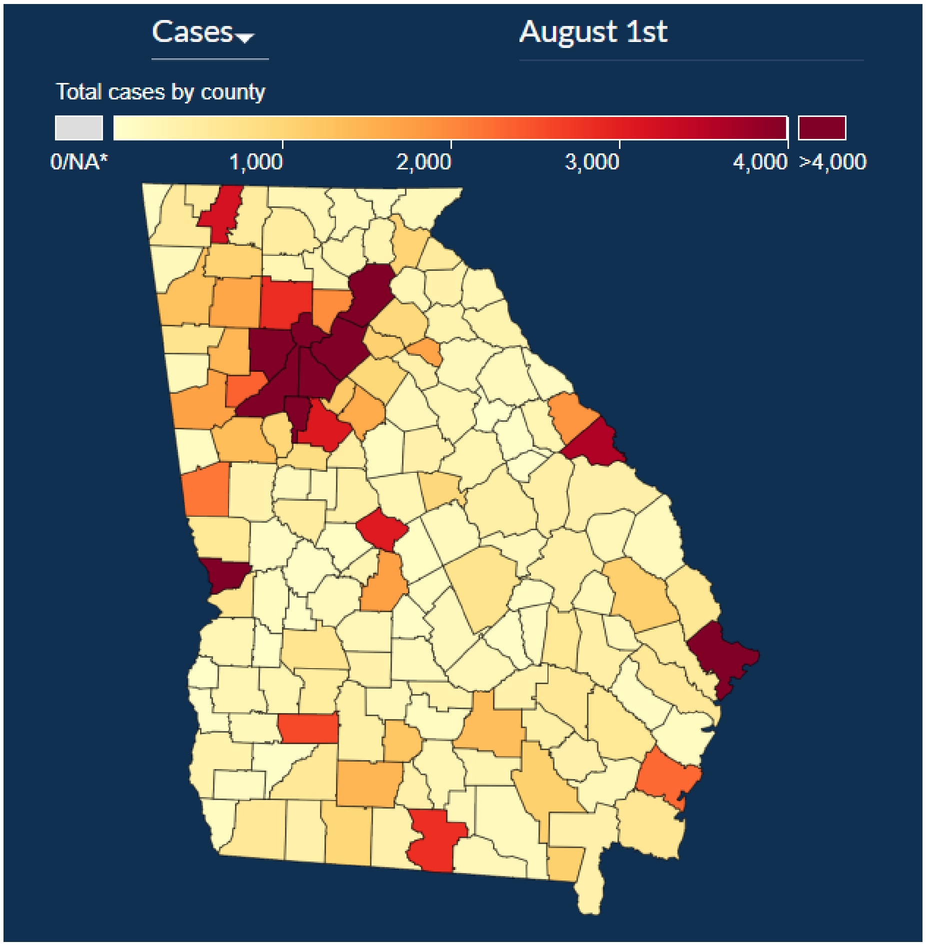



Figure 1.23: Georgia Covid Map

1. **Part I: Using percentage to describe change**

    - In 2020 what is the total number of Gwinnett County residents?

    - Find and record the following information for Gwinnett County.

| Covid 19/Date | April 01-2020 | June 01-2020 | August 01-2020 | Today |
|---|---|---|---|---|
| Cases | | | | |
| Cases Per 100k | | | | |
| Deaths | | | | |

    - For each of the Covid 19 categories of data (Cases, Cases per 100k, Deaths) calculate the **Absolute Change** and **Relative Change** for **April 01 vs June 01** and **June 01 vs August 01** (Remember relative change is a percent change).

- Write a paragraph describing the meaning of these changes, compare the three dates and discuss how the pandemic has been behaving in Gwinnett County.

**Counties** | Top Five Counties | Deaths in Georgia | Race and Ethnicity

| County | ↓ Confirmed Cases | Cases per 100K | Total Deaths | Hospitalizations |
|---|---|---|---|---|
| Fulton | 24891 | 2264.5 | 518 | 2141 |
| Gwinnett | 24252 | 2497.26 | 336 | 2434 |
| Cobb | 16966 | 2146 | 391 | 1589 |
| DeKalb | 16555 | 2087.24 | 300 | 1843 |
| Non-GA Resident/Unknown State | 16233 | 0 | 127 | 561 |
| Hall | 7714 | 3738.33 | 125 | 836 |
| Chatham | 7065 | 2418.06 | 121 | 654 |
| Clayton | 6478 | 2125.06 | 138 | 605 |
| Richmond | 6040 | 2986.55 | 126 | 489 |
| Muscogee | 5466 | 2852.43 | 134 | 579 |

Figure 1.24: The summary of data

2. **Part II: Using percentage as a fraction**

Search for the information about the top five Counties.

- List the top Five counties at the current date
- Using the confirmed cases as a whole, list the percentage of hospitalizations in the top five counties.
- Using the confirmed cases as a whole list the percentage of deaths in the top five counties.
- Use the race and ethnicity information and the confirmed cases as a base. What race or ethnicity has the highest death rate among all.
- With this information, write a paragraph reflecting on what this information means to you. Do you think it is important to have access to this kind of information?
- Go back to the map of Georgia. Do you believe the information you see on the map is displayed properly to understand the numbers?

3. **Part 3: Using Percent as a difference**

Using the categories (Cases, Cases per 100k, Deaths) at the current date, compare Gwinnett County with the other two Counties.

- What Counties did you choose? Find the absolute difference between the Counties.
- Using **April 01 vs August 01** see what county has a larger relative difference.
- Consider the Over time chart for these counties. Are your number reflecting the behavior in the chart? why?

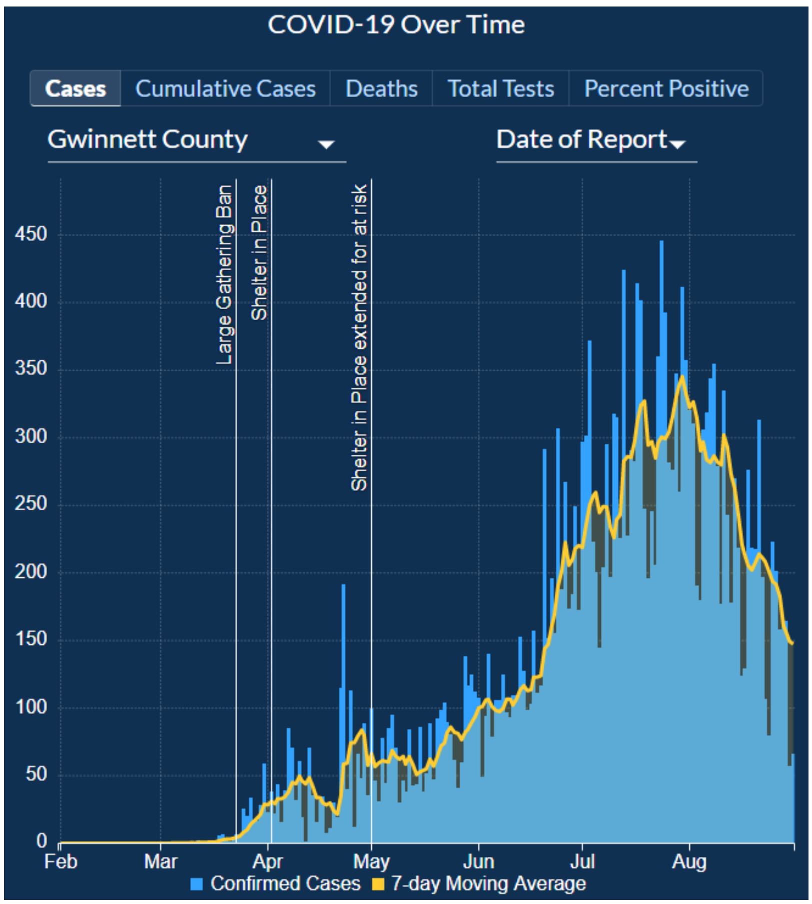
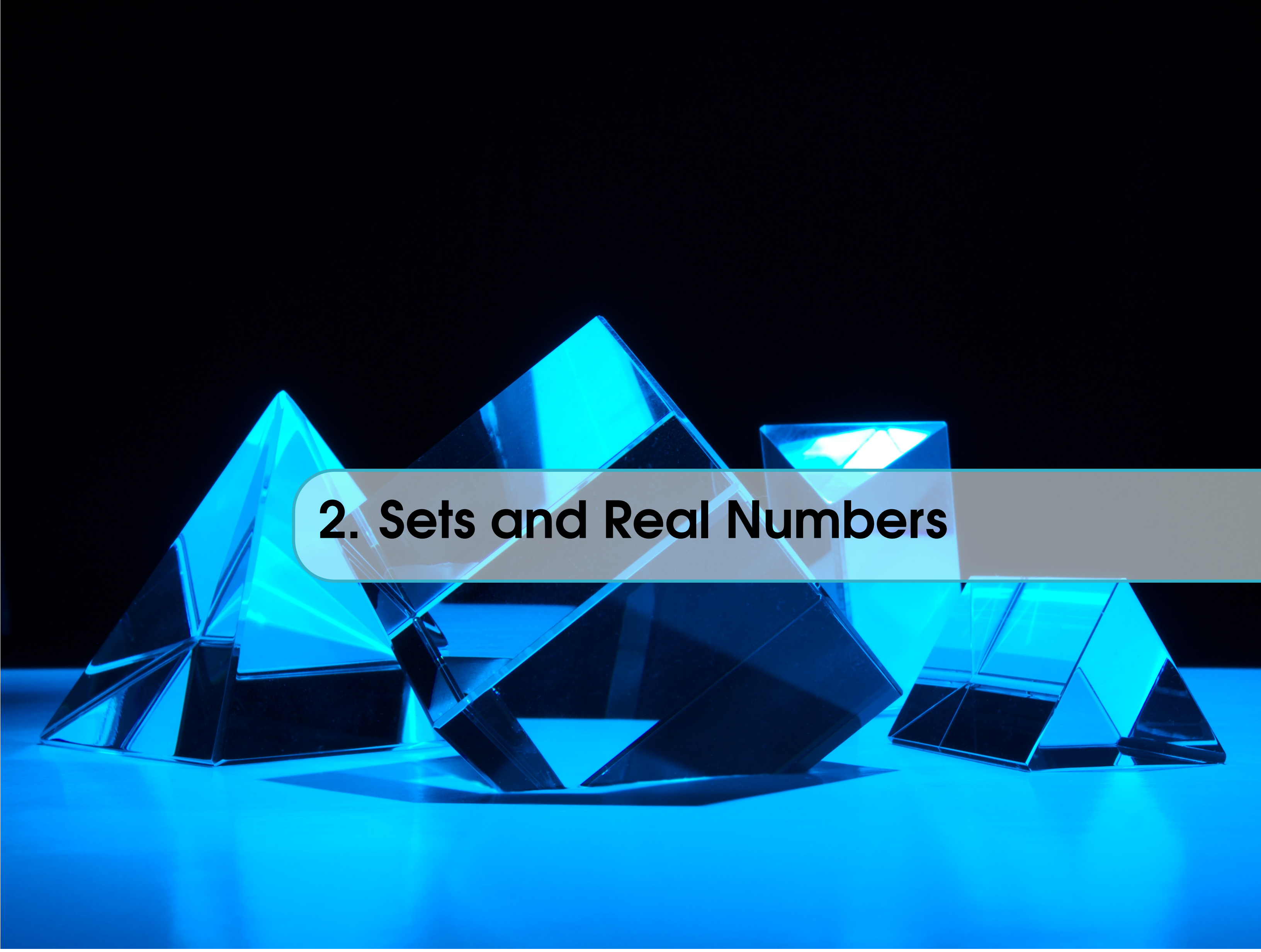



Figure 1.25: Over time figure of Gwinnett in August 01

- What can you conclude from these comparisons?
- Contrast this information with the CDC numbers.



# 2. Sets and Real Numbers

When you think of a set, what do you picture? Is it a set of baseball cards? Or maybe a box of books? Or perhaps a group of your friends? Defining sets in mathematics is almost the same as defining sets in life. A set is just a collection of objects.

In your own major and career, you will find that it will be essential to compare sets. For example, you may be looking at data about the college. You can look at the group of students, the collection of faculty, and the staff set. Do any of these sets have people in common? Is it possible that one of our staff members is currently working on a degree? Then look at the sets of vaccinated and unvaccinated people at the college. How many of the students are in the vaccinated group?

Looking at all of those numbers can give you a headache. But you can organize the numbers into a picture called a Venn Diagram, which will give you a better idea of what you are looking at. In this chapter, we will better define sets and show how we can show sets of numbers with Venn Diagrams.

## 2.1 Sets and Notation

The examples in the chapter opener about a set of baseball cards, a box of books, or your friends fit perfectly with the mathematical definition of a set.

Figure 2.1: A collection of different spices is an excellent example of a set.

**Definition 2.1.1** A **set** is an unordered collection of objects, where each object in the collection is called an **element** of the set. If $x$ is an element of the set $A$, we write

$$x \in A$$

There are two ways to represent a set:

- By list: where we show each element of the set in a list.
- Set Builder Notation: where we construct the set via a property that defines the elements.

Since a set is unordered, it does not matter how the elements are listed. When listing the elements of a set, we use set brackets $\{\}$ surrounding the list of elements of the collection.

**Example 2.1 — Alphabet Sets.**

Consider the following sets:

- $A$ = the set containing the first 5 letters of the English alphabet.
- $B$ = the set containing the vowels in the English alphabet (do not include y as a vowel)
- C = the set of letters in the word "mathematics".

Write each set as a list:

**Solution:**

$A = \{a,b,c,d,e\}$ $\qquad$ $B = \{a,e,i,o,u\}$ $\qquad$ $C = \{m,a,t,h,e,i,c,s\}$

Sometimes, there are too many elements in a set to list them all. For example, if we look at the collection of letters in the English alphabet, there are 26 letters to write. But since they follow an established pattern, we can shorten this by using three dots (...), called an ellipsis, which means "keep going."

■ **Example 2.2 — List Notation.**

Write each set in list notation:

- $D$ = the set of letters in the English alphabet.
- $E$ = the set of counting numbers (positive integers)
- $F$ = the set of negative counting numbers (negative integers)
- $G$= the set of whole numbers between 1 and 5 inclusive (inclusive means including the end numbers, 1 and 5)
- $H$ = the set of whole numbers between 1 and 5 exclusive (exclusive means not including the end numbers)

**Solution**

- $D = \{a, b, c, ..., z\}$
- $E = \{1, 2, 3, ...\}$
- $F = \{..., -3, -2, -1\}$
- $G = \{1, 2, 3, 4, 5\}$
- $H = \{2, 3, 4\}$

■

**Definition 2.1.2** We say that $A = B$ if $A$ and $B$ contain precisely the same elements. We say that A is a subset of $B$, or $A \subset B$ if every element of $A$ is also in $B$.

■ **Example 2.3** Consider the sets

$$A = \{a, b, c, d, e\}, S = \{a, e, d, b, c\}, T = \{a, e, d\}, D = \{a, b, c, ..., z\}$$

What can you say in terms of subsets about A and S? A and T? A and D?

**Solution:**

$A = S$, because they have the same elements. $T \subset A$, because every element of $T$ is also in $A$. $A \subset D$, because every element of $A$ is also in $D$.

Notice we cannot reverse the sets in the second and third examples. A is not a subset of T because $b \in A$ and $e \in A$, but they are not elements of T. So we can write $A \not\subset T$. ■

**Definition 2.1.3** The **empty set** is the set containing nothing. It can be written two ways:

$$\{\} \text{ or } \emptyset.$$

**Definition 2.1.4** The **intersection** of sets A and B, written $A \cap B$, is the set of elements that A and B have in common.

**Definition 2.1.5** The **union** of sets A and B, written $A \cup B$, is the combined set of all elements that are in A or B or both.

**Example 2.4 — American Flag and Rainbow.**

Consider F as the set of colors in the American Flag and R is the set of colors of a Rainbow.

Write the following in the list set notation:

$$F, R, F \cap R, F \cup R$$

**Solution:**

- $F = \{\text{red, white, blue}\}$
- $R = \{\text{red, orange, yellow, green, blue, indigo, violet}\}$
- $F \cap R = \{\text{red, blue}\}$
- $F \cup R = \{\text{red, white, blue, orange, yellow, green, indigo, violet}\}$

■

In the last example, notice that even though red and blue are in both sets when we write the union, $F \cup R$, we only write each word once. There should never be any repetitions in your set.

**Definition 2.1.6** The **Universal Set, U**, for any situation, is everything that could be relevant to that situation.

In the last example, the sets F and R were both sets of colors. So it makes sense to think of the Universal Set, in this case, as the set of all possible colors.

**Definition 2.1.7** The **complement** of a set, A, is the set of elements in the Universal set which are not in A. We write $A^{\mathsf{C}}$ or $A'$ for the complement of set A.

For example, if the Universal set is $U = \{1,2,3,4,5,6,7,8,9,10\}$ and $S = \{1,3,5,7,9\}$, then the complement of S would be $S^{\mathsf{C}} = \{2,4,6,8,10\}$.

■ **Example 2.5** Let

$$U = \{\text{black, red, yellow, orange, blue, white, green, purple}\}$$

and

$$S = \{\text{red, orange, yellow, blue}\}, \quad T = \{\text{black, blue, red}\}$$

Write the following sets in set notation:

(a) $S^{\mathsf{C}}$
(b) $T^{\mathsf{C}}$
(c) $S^{\mathsf{C}} \cap T^{\mathsf{C}}$
(d) $S^{\mathsf{C}} \cup T^{\mathsf{C}}$
(e) $S^{\mathsf{C}} \cap T$
(f) $(S \cup T)^{\mathsf{C}}$.

**Solution:**

(a) $S^{\mathsf{C}} = \{\text{black, green, purple, white}\}$

(b) $T^{\mathsf{C}} = \{\text{ yellow, orange, green, white, purple}\}$

(c) $S^{\mathsf{C}} \cap T^{\mathsf{C}} = \{\text{white, purple}\}$ since $S^{\mathsf{C}}$ and $T^{\mathsf{C}}$ both contain these colors.

(d) $S^{\mathsf{C}} \cup T^{\mathsf{C}} = \{\text{black, green, purple, white, yellow, orange, green}\}$

(e) $S^{\mathsf{C}} \cap T = \{\text{black}\}$ (notice we are looking at $T$ here, not $T^{\mathsf{C}}$)

(f) To find $(S \cup T)^{\mathsf{C}}$, we must first find $S \cup T$ and then take the compliment of that set. Since

$$S \cup T = \{\text{red, yellow, orange, blue, black}\},$$

we can see that

$$(S \cup T)^{\mathsf{C}} = \{\text{green, purple, white}\}.$$

■

### 2.1.1 Number systems as sets

Let's look at the Real Numbers, $\mathbb{R}$, as they were developed throughout history. Early humans only used counting numbers. They did not have a number for zero since they did not know why it would be needed to measure nothing! Counting numbers were the only numbers that were natural to them. Thus our first set of numbers was born: the Natural Numbers.

**Definition 2.1.8** The **Natural Numbers,** $\mathbb{N}$ is the set of counting numbers starting from the number 1.

$$\mathbb{N} = \{1, 2, 3, \ldots\}$$

Notice that 0 is not a Natural Number. When we only used numbers for counting, it did not make sense to count what you do not have. Later, when we developed writing, some ancient numeration systems began to include 0. If you join 0 with the Natural numbers, you get the set of Whole Numbers.

**Definition 2.1.9** The **Whole Numbers,** $\mathbb{W}$ is the set of counting numbers starting from the number 0.

$$\mathbb{W} = \{0, 1, 2, 3, \ldots\}$$

Later in history, scientists needed to talk about moving backward, thus needing numbers less than 0. Bankers also found this helpful in keeping track of money owed, such as in a loan. And in colder places, the temperature outside might fall below zero. Because of this, our negative numbers started coming into play. We added the whole negative numbers to our set of Whole Numbers to get what is called the Integers.

**Definition 2.1.10** The **Integers,** $\mathbb{Z}$, is the set of positive and negative Whole Numbers.

$$\mathbb{Z} = \{\ldots -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

Somewhat before negative numbers came to be, people realized that you could break things into pieces. Before that, if you broke a bone in half, you then had two bones. To talk about these numbers, we needed fractions. So the set of Rational Numbers was born.

**Definition 2.1.11** The **Rational Numbers,** $\mathbb{Q}$ is the set of numbers that can be written as a fraction. $\mathbb{Q}$ is a large set since every integer can be written as a fraction and any decimal that either terminates or repeats.

$$\mathbb{Q} = \left\{ \frac{p}{q} \quad \text{such that} \quad p \in \mathbb{Z} \quad \text{and} \quad q \in \mathbb{N} \right\}$$

$\mathbb{Q}$ is the set of numbers that can be written as a fraction with the numerator being an integer and the denominator being a natural number. Some numbers cannot be written as a fraction.

$\sqrt{2}$ and $\pi$ are examples of these numbers. Such numbers are known as Irrational Numbers.

**Definition 2.1.12** The **Irrational Numbers,** , is the set of numbers that cannot be written as fractions. These include all numbers with decimals that do not end and never form a repeating pattern.

$$\mathbb{I} = \{\text{real numbers that are not rational}\}$$

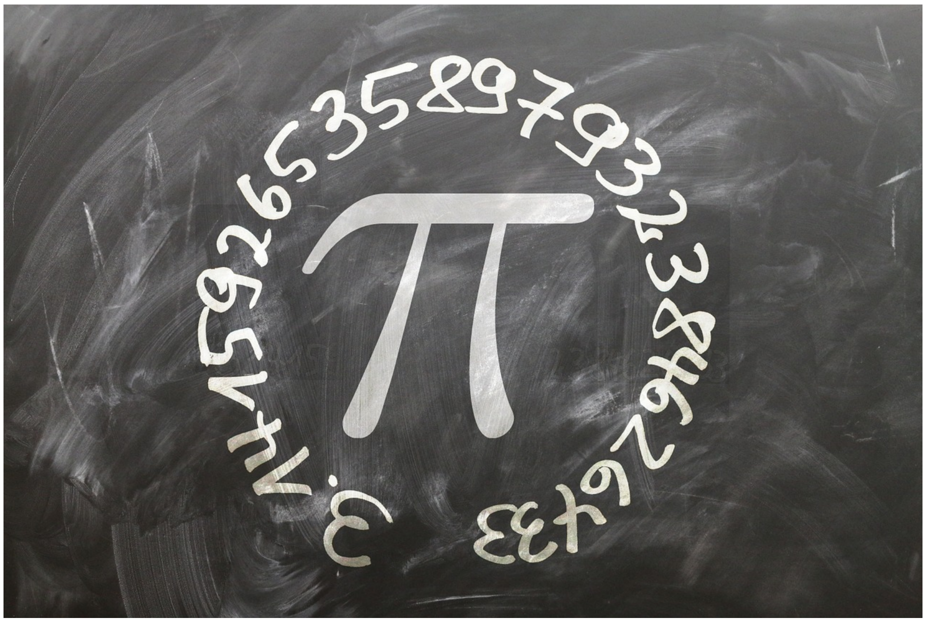


Figure 2.2: One of the most famous irrational numbers: Pi.

All of the numbers we have talked about make up the set of Real Numbers, $\mathbb{R}$.

■ **Example 2.6** Tell in which sets each number belongs, chosen from $\mathbb{N}, \mathbb{W}, \mathbb{Z}, \mathbb{Q}, \mathbb{I}, \mathbb{R}$:

(a) 3
(b) 0
(c) $\frac{5}{8}$
(d) $-\frac{1}{2}$
(e) $\sqrt{5}$
(f) 0.34
(g) 0
(h) $0.\overline{6}$

**Solution:**

(a) $3 \in \mathbb{N}, \mathbb{W}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

(b) $0 \in \mathbb{W}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

(c) $\frac{5}{8} \in \mathbb{Q}, \mathbb{R}$

(d) $-4 \in \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

(e) $-1/2 \in \mathbb{Q}, \mathbb{R}$

(f) $\sqrt{5} \in \mathbb{I}, \mathbb{R}$

(g) $0.34 \in \mathbb{Q}, \mathbb{R}$

(h) $0.\overline{6} \in \mathbb{Q}, \mathbb{R}$

The hardest part of this example is deciding whether or not a number is a Rational Number. The reasoning for some of the answers follows:

$$3 = \frac{3}{1}, 0 = \frac{0}{1}, -4 = \frac{-4}{1}, 0.34 = \frac{34}{100},$$

and

$$0.\overline{6} = 0.66666... = \frac{2}{3}$$

■

Looking back at the previous example, you should notice that every Natural Number is a Whole Number, every Whole Number is an Integer, every integer is a Rational Number, and every Rational Number is a Real Number. In fact,

$$\mathbb{N} \subset \mathbb{W} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

Also, every Irrational Number is a Real Number, so $\mathbb{I} \subset \mathbb{R}$ as well.

Figure 2.3 shows the make-up of the Real Number System and how each of the sets we have discussed is related to the others. Notice: None of the other subsets of $\mathbb{R}$ that we have mentioned are in the set of Irrational numbers.

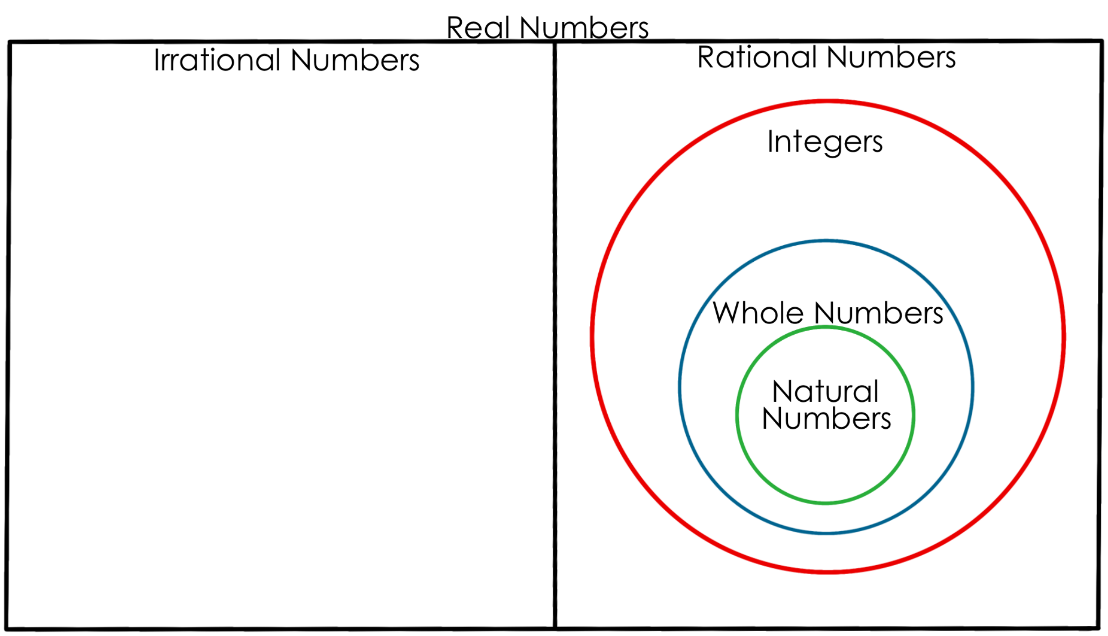



Figure 2.3: Venn diagram showing how the different subsets of the Real Numbers are associated. Notice that after you split the Irrational Numbers away from the Rational Numbers, that side of the diagram does not have any of our other subsets inside it.

In this diagram, the universe could be considered the set of the Real Numbers, with the Rational and Irrational Numbers complementing each other.

### 2.1.2 Exercises Section 2.1

**Problem 2.1** Use set notation to list the elements of each set.

(a) The set of positive even numbers less than 10

(b) The set of positive even numbers

(c) The set of whole numbers from 1 to 1000 inclusive.

(d) The set of odd integers.

(e) The set of even integers bigger than zero and less than 12.

**Problem 2.2** If $U = \{1,2,3,\ldots,20\}$, $A = \{2,4,6,8,10\}$, $B = \{5,10,15,20\}$, and $C = \{3,6,9,12,15\}$ write the following sets in set notation.

(a) $A \cap B$

(b) $A \cap C$

(c) $B \cup C$

(d) $A \cap B'$

(e) $A \cap B \cap C$

(f) $C' \cup B$

(g) $C'$

(h) $(A \cup B)'$

(i) $A' \cap B'$

**Problem 2.3** Write three different numbers that are elements of the set.

(a) $\mathbb{Z}$ but not $\mathbb{N}$.

(b) $\mathbb{I}$

(c) $\mathbb{Z}$ and $\mathbb{Q}$

(d) $\mathbb{R}$

**Problem 2.4** Write three different numbers that are elements of the set.

(a) $\mathbb{Z}$ but not $\mathbb{N}$.

(b) $\mathbb{I}$

(c) $\mathbb{Z}$ and $\mathbb{Q}$

**Problem 2.5** Let

$$R = \{\text{apple, orange, cherry, lemon, mango}\} \quad S = \{\text{lemon, orange}\}$$

$$T = \{\text{lime, lemon, orange, grapefruit}\} \quad V = \{\text{orange, lemon}\}$$

State whether each statement below is true or false.

(a) $R \subset S$

(b) $S \subset R$

(c) $S \subset T$

(d) $S \subset T \cap R$

(e) $R = T$

(f) $S = V$

(g) $S = V'$

(h) $V \subset R'$

**Problem 2.6** Using the same $R, T, S, V$ as in the previous problem. Calculate:

(a) $R \cap V$

(b) $T \cup S$

**Problem 2.7** For each number listed below, decide which of these sets contain it: $\mathbb{N}, \mathbb{W}, \mathbb{Z}, \mathbb{Q}, \mathbb{I}, \mathbb{R}$.

(a) $\frac{4}{3}$

(b) $5.\overline{238}$

(c) 1683

(d) 0

(e) $\pi$

(f) $-25$

(g) $-31.4$

(h) $\frac{10}{5}$

**Problem 2.8** State whether each of the following statements is True or False.

(a) There are numbers in $\mathbb{Q}$ that are also in $\mathbb{W}$

(b) Negative Integers are not Real Numbers

(c) $\mathbb{Q}$ is a subset of $\mathbb{I}$

(d) $\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}$

(e) $\mathbb{Z} \subset \mathbb{N}$

(f) $\mathbb{W} = \mathbb{N} \cup \mathbb{R}$

**Problem 2.9** Define each term below:

(a) Complement

(b) Element

(c) Empty Set

(d) Integer

(e) Intersection

(f) Irrational Number

(g) Natural Number

(h) Rational Number

(i) Real Number

(j) Set

(k)Subset

(l) Union

(m) Universal Set

(n) List notation

**Problem 2.10** Write each set using set-list notation.

(a) The set of integers from 8 to 12 inclusive.

(b) The set of perfect squares that are less than 100.

(c) The set of odd whole numbers.

(d) The set of whole numbers between 1 and 100 inclusive.

(e) A set with 5 irrational numbers.

**Problem 2.11** Tell in which sets each number belongs, chosen from $\mathbb{N}, \mathbb{W}, \mathbb{Z}, \mathbb{Q}, \mathbb{I}, \mathbb{R}$.

(a) 0

(b) $\sqrt{28}$

(c) $-100$

(d) $\frac{32}{71}$

(e) 1

(f) 275

(g) $2\pi$

(h) $3.\overline{2}$

(i) $\frac{1}{2}$

(j) $-13$

(k) $e$

(l) $\sqrt{2}$

**Problem 2.12** Let the Universal set $U = \{a, b, c, \ldots, z\}$ and

$A = \{a, b, c, d, e\}$

$B = \{a, b, c, \ldots, m\}$

$C = \{r, h, o, m, b, u, s\}$

$D = \{w, x, y, z\}$

Write the following sets using set notation.

(a) $A \cap B$

(b) $B^{\complement}$

(c) $C \cap D$

(d) $C \cup D$

(e) $A \cup D$

**Problem 2.13** Let the Universal set $U = \{a,b,c,...,z\}$ and

$A = \{a,b,c,d,e\}$

$B = \{a,b,c,...,m\}$

$C = \{r,h,o,m,b,u,s\}$

$D = \{w,x,y,z\}$

Decide if each statement is True or False.

(a) $B \subset A$

(b) $A \subset B$

(c) $A \cup B = D$

(d) $A \cup D \subset U$

(e) $A \cap C \subset C$

**Problem 2.14** Write the following sets in set builder notation ( Describing the sets by a property).

(a) $A = \{2,4,6,8,10,12\}$

(b) $B = \{1,3,5,7,9,11\}$

(c) $C = \{-1,-3,-5,-7\}$

(d) $D = \{1,2,3,4,5\}$

(e) $E = \{2,3,5,7,11,13,17,19\}$

**Problem 2.15** Assuming we have a universal set given by:

$$U = \{x \in \mathbb{N} : 1 \leq x \leq 20\}$$

and

$A = \{4,6,8,10,12,15,20\}$

$B = \{3,5,7,9\}$

$C = \{1,2,3,4,5\}$

$D = \{2,3,5,7,11,13,17,19\}$

Find

(a) $A^{\complement}$ (b) $B^{\complement}$ (c) $C^{\complement}$ (d) $D^{\complement}$

## 2.2 Sets and Venn Diagrams

Now that we have learned the ideas of a set, a union, an intersection, a subset, and the empty set, it would be nice to have a way to picture our sets. In Figure 2.3, you saw a picture of the real numbers. This type of picture is called a Venn Diagram.

A Venn Diagram comprises a rectangle, with circles or other shapes showing how sets are related. Let's start with a simple example.

Let $A = \{a,b,c,d,e,f,g,h,i,j\}$ and $V = \{a,e,i,o,u\}$ . Figure 2.4 shows a Venn diagram illustrating these sets. Notice that since $A \cap V = \{a,e,i\}$, those letters have been placed in the overlapping section of the two circles. The letters in A that are not vowels are in the left part of A so that they don't touch V. Similarly, o and u are vowels but not part of A, so they are in the right part of V.

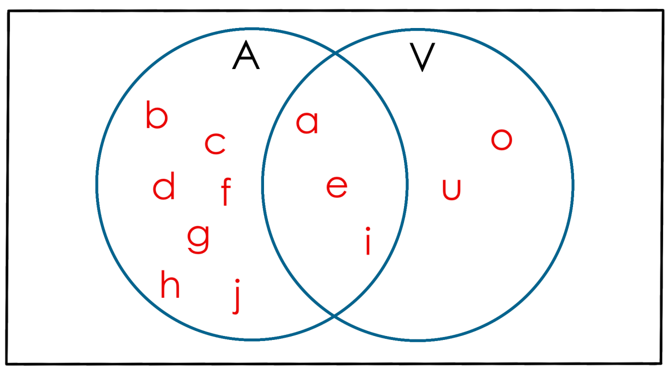
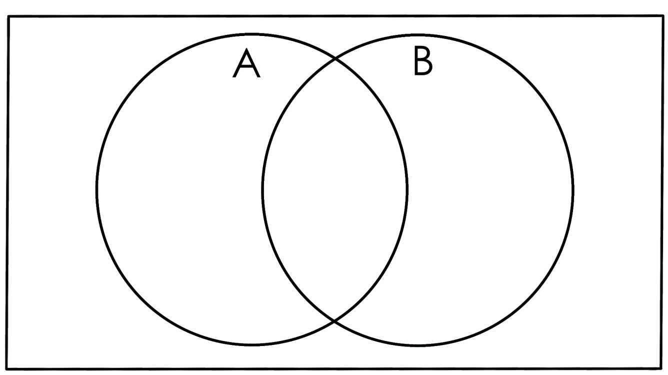


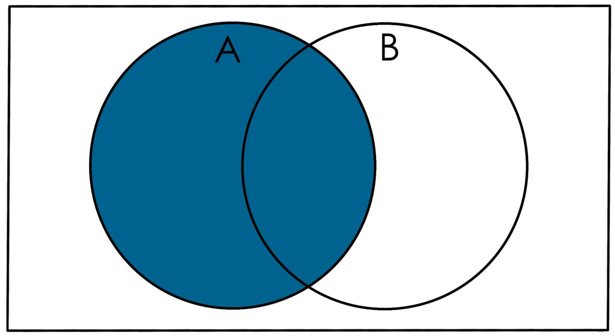

Figure 2.4: This Venn diagram shows A and V, but includes the letters outside of both.

■ **Example 2.7** Using the Venn diagram shown below, shade the corresponding part for the following sets: $A, A \cap B, A \cup B, A'$



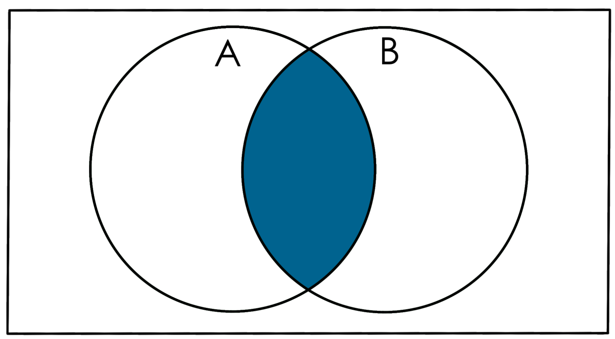

**Solution:**

- $A$



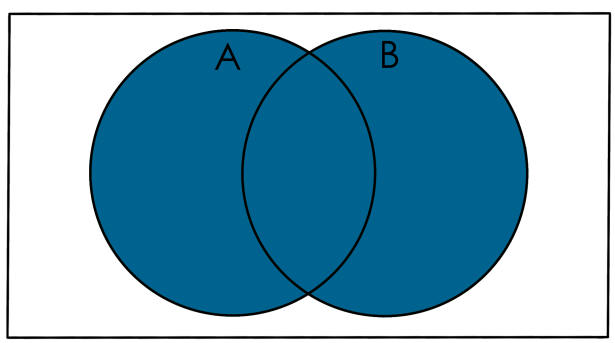

Here notice only the elements corresponding to $A$ are shaded.

- $A \cap B$



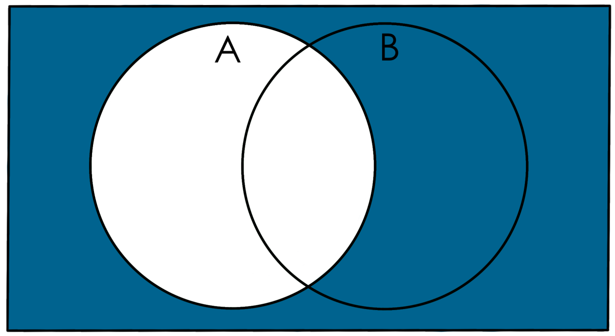

In this case, we shaded the common elements between $A$ and $B$.

- $A \cup B$




The shaded region for the union includes the elements both in $A$ or $B$.

- $A'$




∎

Notice how we shade everything outside of the particular set we are drawing complement for.

In Figure 2.4, we ignored the letters of the alphabet that are not included in A or V. For completeness, let's assume that the English alphabet is the Universal set for this case. Then Figure 2.5 would show the entire situation.

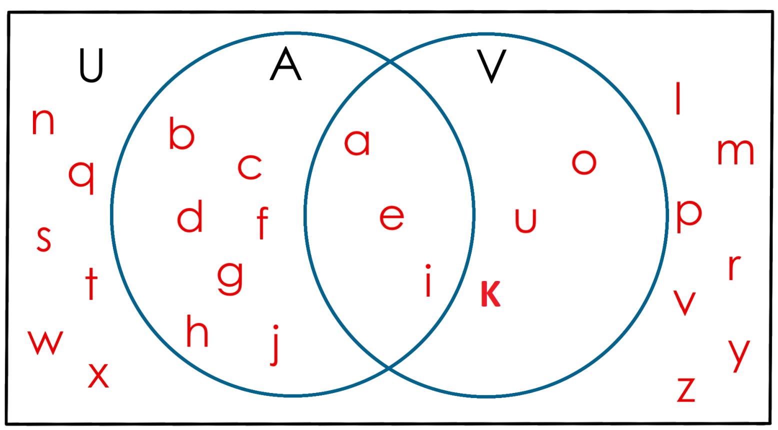



Figure 2.5: This Venn diagram shows A and V, but includes every other letter of the English alphabet.

Now that we know the basics of Venn diagrams let's see how two use them to count when two or three sets are involved.

### 2.2.1 Counting with 2- and 3- circle Venn Diagrams

It is often important to know how many objects are in a set. Let's say that $D = \{$dog, cat, bird, lizard$\}$. Then the number of elements of $D$, or $n(D) = 4$. When we have more than one set, you want to know how many objects are in each set and how many they have in common.

For example if $D = \{$dog, cat, bird, lizard$\}$ and $F = \{$wolf, dog, coyote$\}$ you can count that $n(D) = 4$ and $n(F) = 3$. But what about $n(D \cap F)$ ? Since $D \cap F = \{$dog$\}$,we can see that $n(D \cap F) = 1$. We could also Find $n(D \cup F)$. Since $D \cup F =$dog, cat bird, lizard, wolf, coyote, then $n(D \cup F) = 6$.

Let's look at the two Venn diagrams in Figure 2.6 and 2.7.

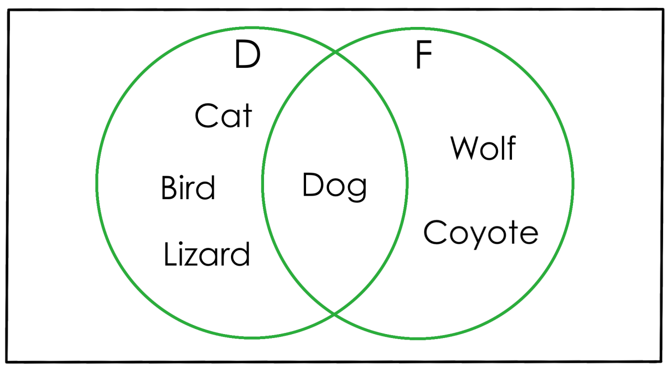



Figure 2.6: This Venn diagram shows what elements are in the set.

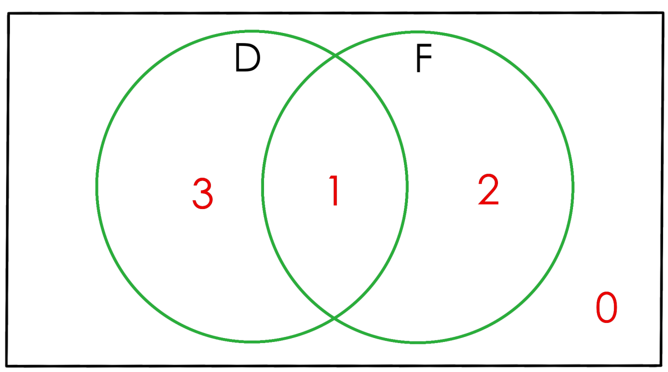
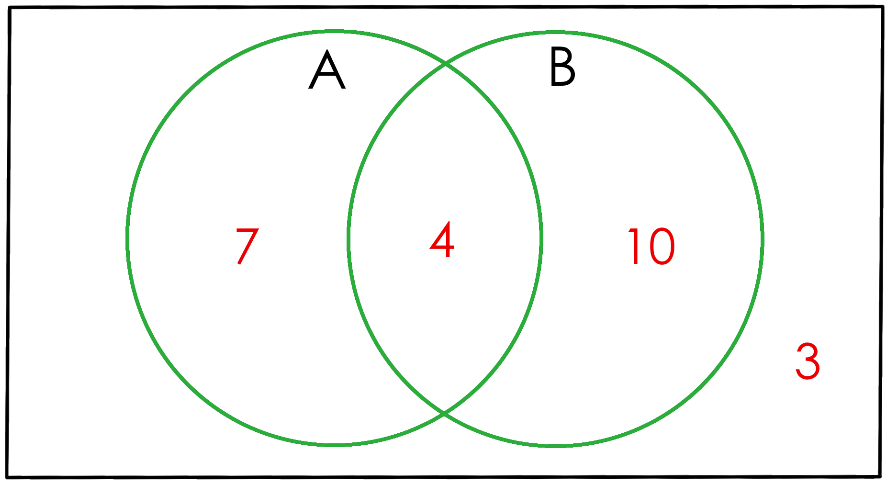



Figure 2.7: This Venn diagram shows how many elements each part of the diagram has, but both diagrams represent the same set.

Remember that we said that $n(D) = 4$. If you look at the circle $D$ in the picture on the right, it contains two numbers: 3 and 1. If we add these, we get a total of 4. So our picture also shows that there are four elements in $D$.

■ **Example 2.8** Find the number of elements in Each set as shown in the Venn Diagram.




(a) $n(A)$

(b) $n(B)$

(c) $n(A^{\complement})$

(d)$n(B^{\complement})$

(e). $n(A \cup B)$

(f)$n(A \cap B)$

**Solution:**

(a) $n(A) = 7 + 4 = 11$

(b) $n(B) = 10 + 4 = 14$

(c) $n(A^{\complement}) = 10 + 3 = 13$

(d) $n(B^{\complement}) = 7 + 3 = 10$

(e) $n(A \cup B) = 7 + 4 + 10 = 21$

(f) $n(A \cap B) = 4$

■

Notice that when we count for sets like the union is essential not to measure twice for the intersection of sets.

If you are looking at three sets, you will have more sections to count. It is important to know what each section means. Look at Figure 2.8. Notice that there are eight regions in this Venn diagram. If you count them and only see 7, you probably forgot about the outer part of the diagram.

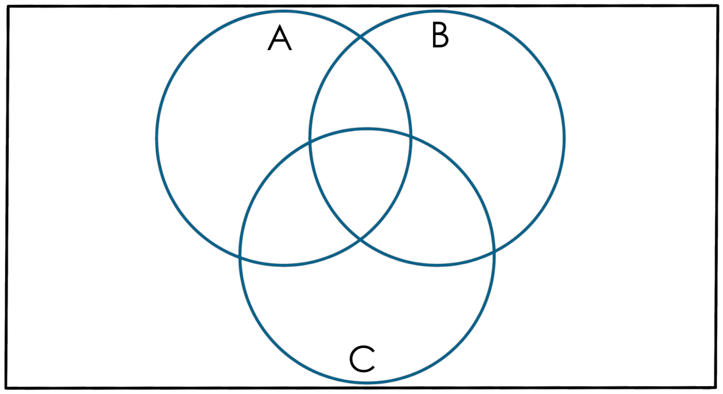



Figure 2.8: A Venn diagram with three circles.

■ **Example 2.9** In the following sets, which section(s) should be shaded in a three-circle Venn diagram?

(a) $C$

(b) $A \cap C$

(c) $B \cup C$

(d) $A \cap B \cap C$

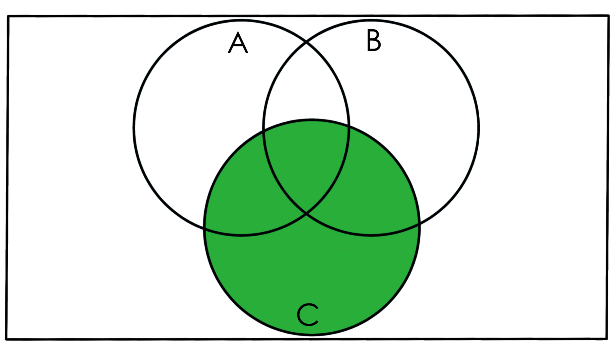
(e) $A \cap B^{\mathsf{C}}$

(f) $(A \cap B)^{\mathsf{C}}$

(g) $(A \cup B \cup C)^{\mathsf{C}}$

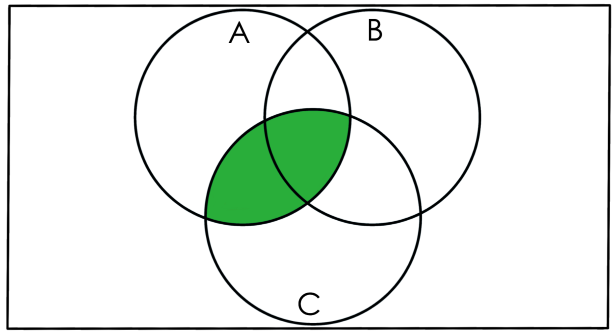
**Solution:**

(a)



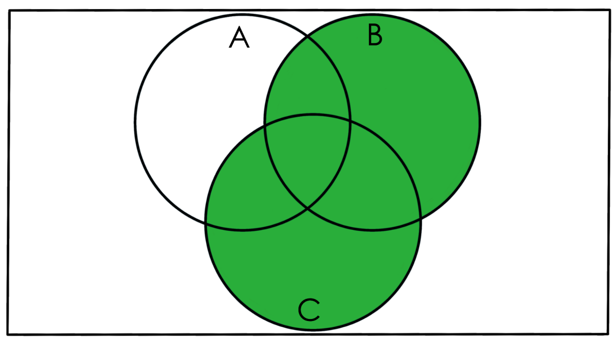

This diagram shows how only the circle related to $C$ is shaded.

(b)



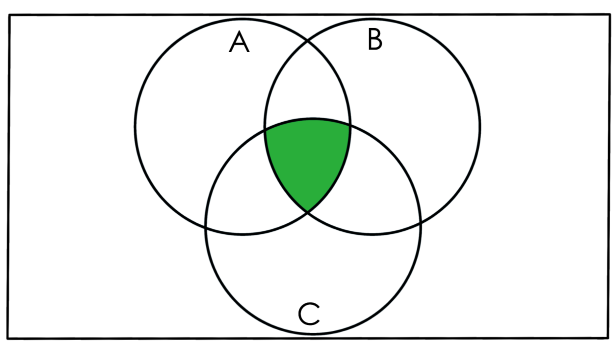

This diagram highlights the common elements between $A$ and $C$.

(c)



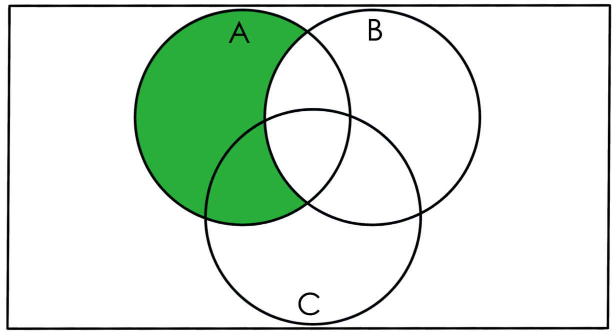

This diagram highlights All elements in $B$ together with $C$.

(d)



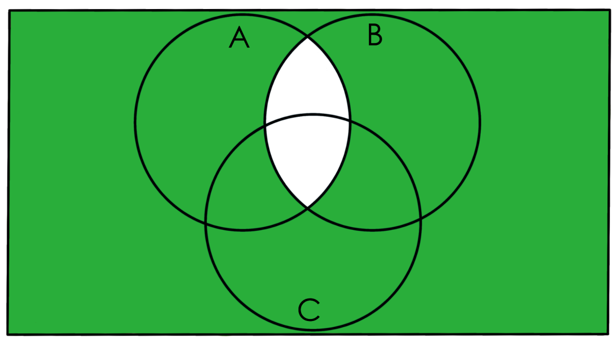

This diagram highlights All the common elements for the three sets.

(e)



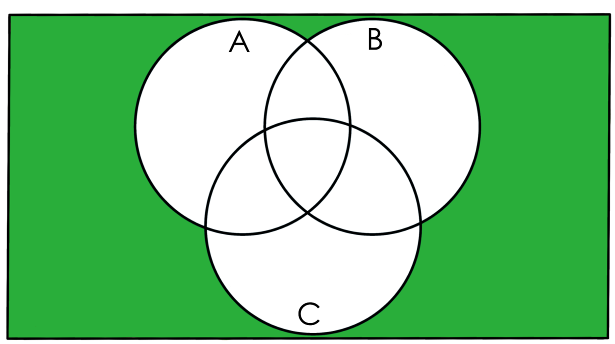

This diagram highlights the elements in $A$ that are not in other sets.

(f)



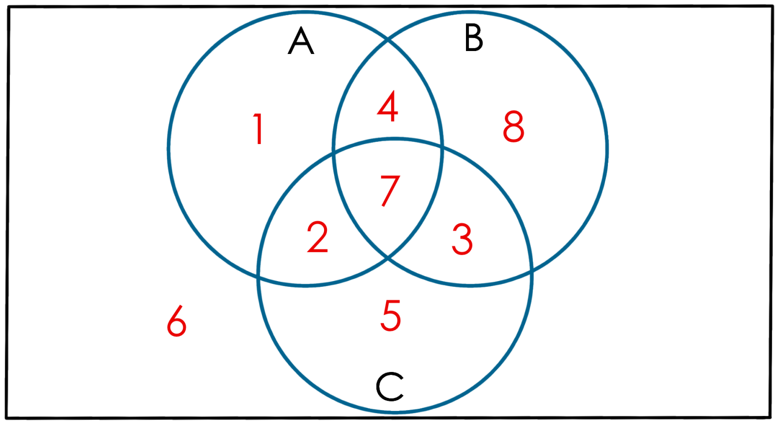

This diagram highlights the elements outside the intersection of $A$ and $B$.

(g)




This diagram highlights the elements outside the three sets. ■

■ **Example 2.10** Using the three circle Venn diagram:




Find:

(a) $n(B)$

(b) $n(A \cap B)$

(c) $n(A \cup C)$

(d) $n(C^{\complement})$

(e) $n(B \cup A^{\complement})$

(f) $n(A \cap B \cap C)$

(g) $n(A \cup (B \cap C))$

**Solution**

(a) $n(B) = 4+8+7+3 = 22$

(b) $n(A \cap B) = 4+7 = 11$

(c) $n(A \cup C) = 1+4+2+7+5+3 = 22$

(d) $n(C^{\complement}) = 1+4+8+6 = 19$

(e) $n(B \cup A^{\complement}) = 4+8+7+3+5+6 = 33$

(f) $n(A \cap B \cap C) = 7$

(g) $n(A \cup (B \cap C)) = 1+4+2+7+3 = 17$

■

As we can see from the respective sums, it is essential not to repeat counting in the intersections, especially when we may have multiple intersections between the three sets.

### 2.2.2 Survey problems

Previously we used Venn diagrams to count how many elements are in different parts of sets. We will now take that one step further and look at Venn diagrams showing answers to surveys. The only new idea is for you to realize that:

(1). When using the word AND, you are taking the intersection of two sets.

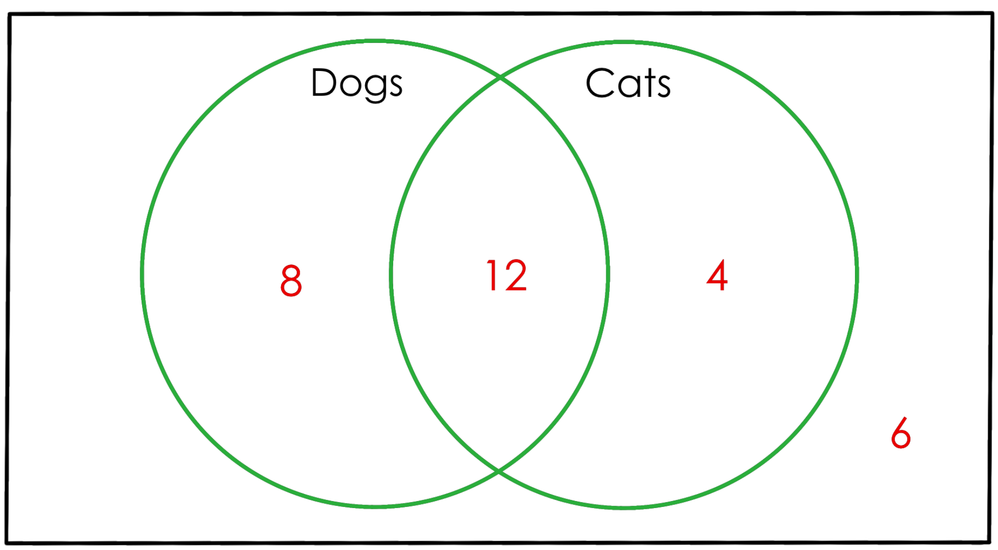
In other words, if I say I like apples and I like oranges, that is the same as saying that I am at the intersection of the set people who like apples and the set of people who want oranges.

(2). When using the word OR, you take the union of two sets.

If I say it is cloudy or hot outside, that means today is in the union of the set of cloudy days and the set of hot days.

In other words: $AND = \cap$ and $OR = \cup$.

■ **Example 2.11** Students were surveyed and asked if they like cats and if they like dogs. The results are in the following Venn diagram:




Answer the following questions:

(a) How many students were surveyed?

(b) How many students like dogs?

(c) How many students like cats?

(d) How many students like both cats and dogs?

(e) How many students like cats or dogs?

(f) How many students like dogs, but not cats?

(g) How many students don't like either of them?

(h) How many students like cats, but not dogs?

(i) How many students do not like cats?

(j) How many students do not like dogs?

**Solution:**

(a) $8+12+4+6=30$ students were surveyed.

(b) $8+12=20$ students like dogs.

(c) $12+4=16$ students like cats.

(d) 12 students (the intersection) like both cats and dogs.

(e) $8+12+4=24$ students (the union) like cats or dogs.

(f) 8 students like dogs, but not cats.

(g) 6 students don't like any of them.

(h) 4 students like cats, but not dogs.

(i) $8+6=14$ students do not like cats.

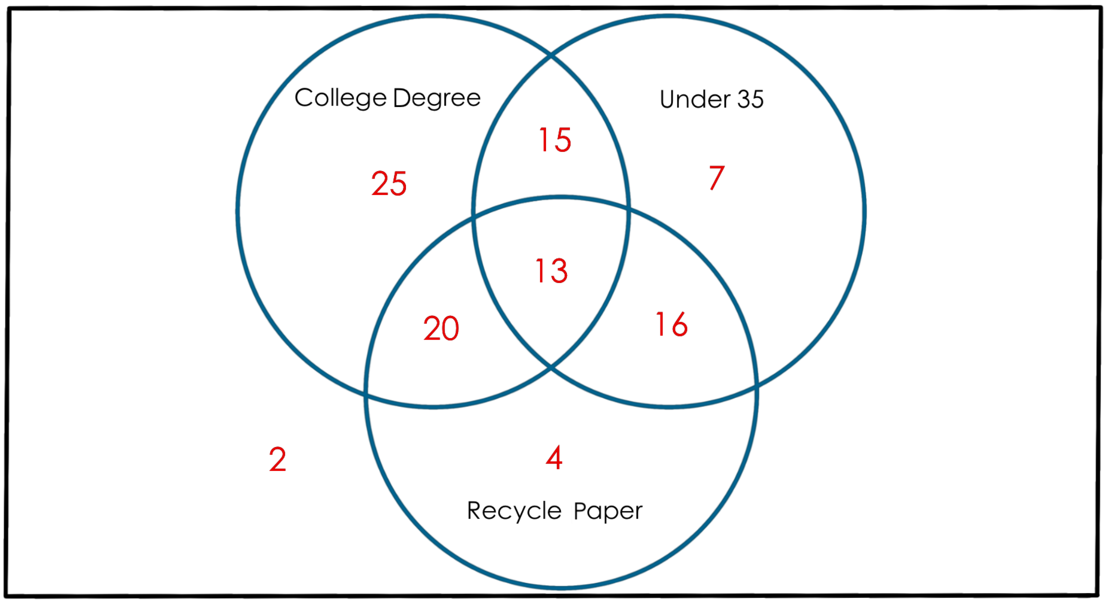
(j) $4+6=10$ students do not like dogs.

■

■ **Example 2.12** Business owners were asked three questions:

- Do you have a college degree?
- Are you under 35 years of age?
- Does your business recycle paper?

The results are in the following diagram:




Answer the following questions.

(a) How many people had a college degree?

(b) How many people had a college degree and were under 35 years of age?

(c) How many people recycled paper, but did not have a college degree?

(d) How many people were not under 35?

(e) How many people were under 35 or had a college degree?

(f) How many poeple were under 35, but did not have a college degree and did not recycle paper?

(g) How many people were under 35, recycled paper, and had a college degree?

(h) How many people did not have a college degree, did not recycle paper and were not under 35?

(i) How many people had a college degree or were under 35 or recycled paper?

(j) How many people were surveyed?

(k) How many people were under 35 and recycled paper?

(l) How many people recycled paper or had a college degree, but were not under 35?

(m) How many people were under 35 and recycled paper, but did not have a college degree?

**Solution:**

(a) $25+15+20+13=73$ people had college degrees.

(b) $15+13=28$ people had a college degree and were under 35.

(c) $4+16=20$ people recycled paper, but did not have a college degree.

(d) $25+20+4+2=51$ people were not under 35.

(e) $25+15+7+20+13+16=96$ people were under 35 or had a college degree.

(f) 7 people were only in the under 35 circle.

(g) 13 people were in all three circles.

(h) 2 people were in none of the three circles.

(i) $25+15+7+20+13+16+4=100$ people were in the union of the three circles.

(j) $25+15+7+20+13+16+4+2=102$ people were surveyed.

(k) $13+16=29$ people were under 35 and recycled paper.

(l) $25+20+4=49$ people recycled paper or had a college degree, but were not in the under 35 circle.

(m) 16 people were under 35 and recycled paper, but did not have college degree.

■

### 2.2.3 Exercises Section 2.2

**Problem 2.16** Draw a Venn Diagram for overlapping sets $A$ and $B$. Then shade the following regions.

(a) $A$

(b) $A'$

(c) $A \cup B$

(d) $A \cap B$

(e) $A' \cap B'$

(f) $A' \cap B$

(g) $A' \cup B'$

(h) $(A \cup B)'$

(i) $(A \cap B)'$

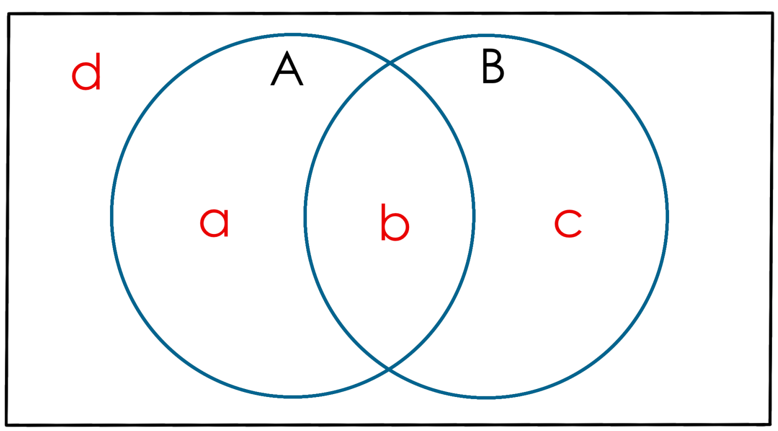
(j) $A \cap B'$

(k) $A \cup B'$

(l) $(A' \cap B')$

**Problem 2.17** Look back at your Venn diagrams e, g, h, and k from the previous problem. Compare these pictures. What do you notice about them?

**Problem 2.18** Match each letter (The red a, b, c or d) from the diagram to the equivalent set listed below.



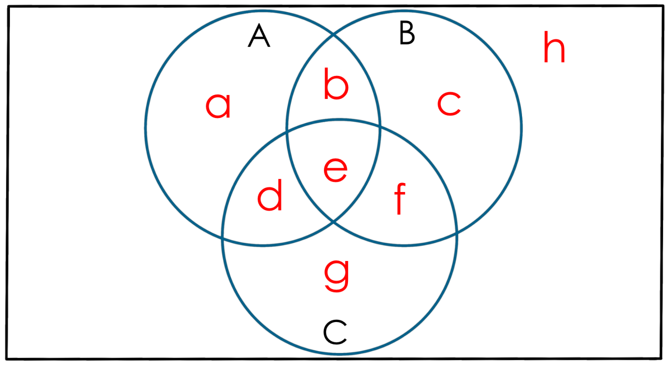

(i) $A \cap B$

(ii) $A \cap B'$

(iii) $(A \cup B)'$

(iv) $A' \cap B$

**Problem 2.19** Match the zones marked with red letters with the equivalent set listed.




(i) $A \cap B \cap C$

(ii) $C \cap (AUB)'$

(iii) $(A \cup B \cup C)'$

(iv) $A \cap (B \cup C)'$

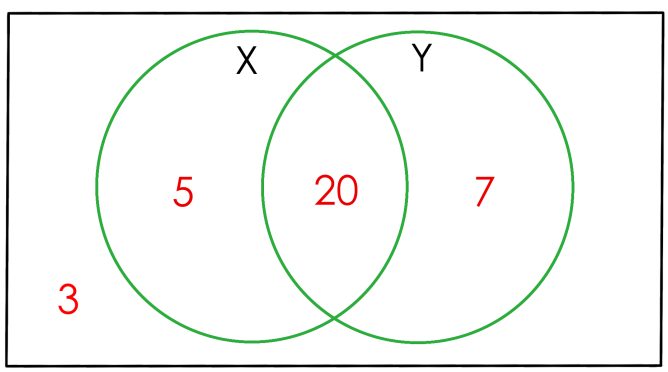
(v) $A \cap B \cap C'$

(vi) $C \cap B \cap A'$

(vii) $B \cap (A \cup B)'$

(viii) $A \cap C \cap B'$

**Problem 2.20** How many elements are in each set listed below.




(a) $X \cap Y$

(b) $X \cup Y$

(c) $(X \cup Y)'$

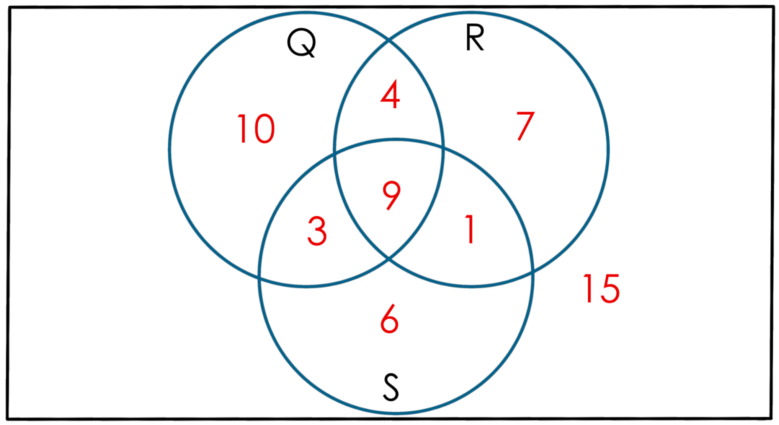
(d) $X' \cap Y$

(e) $(X \cap Y)'$

(f) $X' \cap Y'$

(g) $X' \cup Y'$

**Problem 2.21** How many elements are in each set listed below.




(i) $Q$

(ii) $Q'$

(iii) $Q \cap R$

(iv) $Q \cap R \cap S'$

(v) $Q \cup R \cup S$

(vi) $(S \cap R) \cup Q$

(vii) $R \cap R'$

(viii) $S \cup S'$

(ix) $R$

(x) $Q \cap R \cap S$

**Problem 2.22** For each question, draw a Venn Diagram for overlapping sets A and B. Then shade the set indicated.

(a) $A \cup B$

(b) $A \cap B$

(c) $A'$

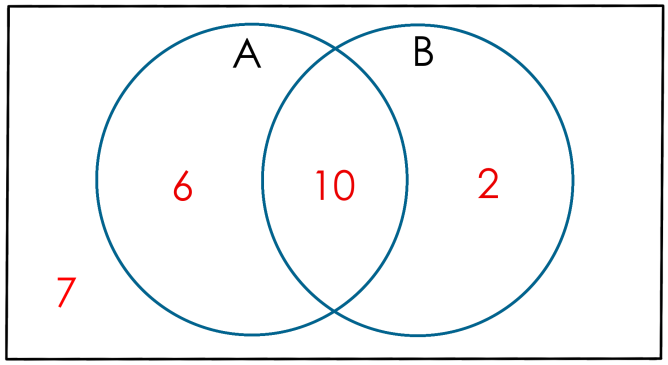
(d) $B'$

(e) $A$

(f) $A \cap B'$

(g) $(A \cup B)'$

**Problem 2.23** Use the Venn diagram to answer the following questions.




(a) $n(A)$

(b) $n(A')$

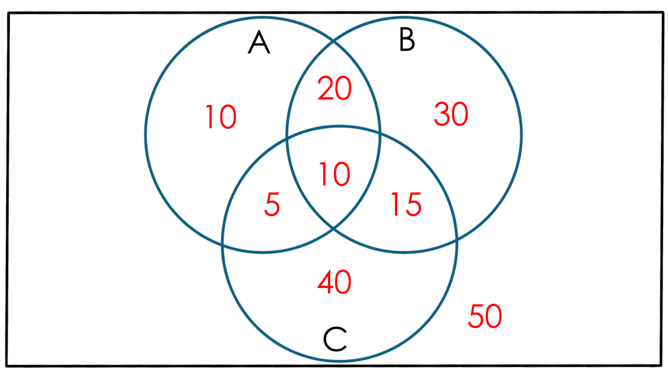
(c) $n(B \cap A)$

(d) $n(B \cup A')$

(e) $n(A \cup B)$

(f) $n(B \cap A)$

**Problem 2.24** Use the Venn diagram to answer the following questions.




(a) $n(A)$

(b) $n(B \cap C)$

(c) $n(A \cap B \cap C)$

(d) $n(C')$

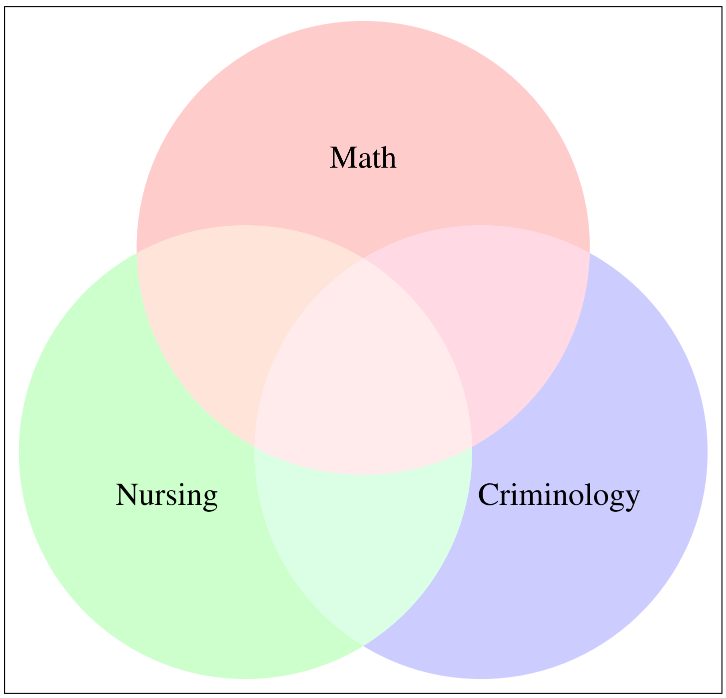
(e) $n(A \cup C)$

(f) $n[(A \cup B \cup C)']$

(g) $n(A' \cap C')$

**Problem 2.25** We interviewed 100 students about their class majors selection in the semester. From them we collected the following information:

- 6 took Criminology and Nursing
- 28 took Math classes
- 31 took Criminology classes
- 9 took Math and Criminology classes
- 10 took Math and Nursing classes
- 42 took Nursing classes
- 4 took classes in all three majors




(a) How many took none of the three subjects?

(b) How many took Math but not Nursing or Criminology?

(c) How Many took Criminology but not Math or Nursing?

(d) How many took Math classes exclusively?

(e) How many took Nursing but not Math?

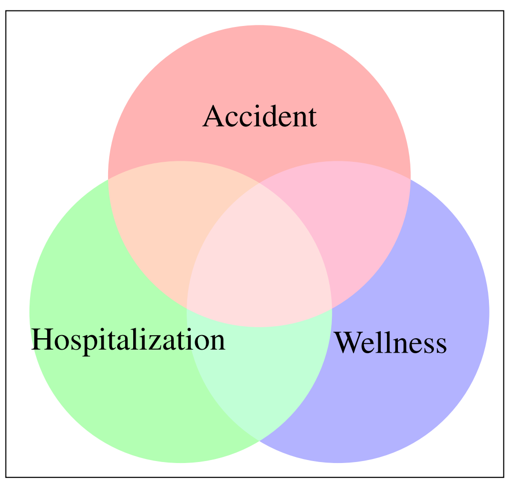
**Problem 2.26** A study records the preferences in beverages between tea and coffee for a groups of men and women with a table as follows:

| | Female | Male | Total |
|---|---|---|---|
| Tea | 22 | | 40 |
| Coffee | | 15 | |
| Total | 51 | 33 | |

- Fill the missing info in the table.
- Using a Venn diagram find how many in the study were male and preferred Coffee.
- With the same Venn diagram find how many were female and preferred Tea.

**Problem 2.27** A hospital receives a list of 200 patients in a day, your job is to classify the kind of insurance each patient has for billing purposes. You have that:

- 30 of them have no insurance.
- 120 have Wellness insurance.
- 10 have only Wellness and Hospitalization (no Accident insurance).
- 40 have Accident and Hospitalization insurance.
- 30 have Comprehensive care insurance (All types)
- 5 have Wellness and Accident insurance (no Hospitalization).
- 20 have only Accident insurance.




(a) How many have wellness insurance only?

(b) How many is the total number of people insured in accidents?

(c) How many have hospitalization insurance only?

## 2.3 Project

### 2.3.1 Project: Take a Survey

This project aims to create a survey that will allow you to fill correctly in a 3-circle Venn diagram showing all of the information.

**Instructions:**

(1). Write a three-question survey in which you only use yes or no questions.
(2). Survey at least 50 people, and have them answer all of the questions.
(3). Take the data and use it to fill in a Venn diagram.

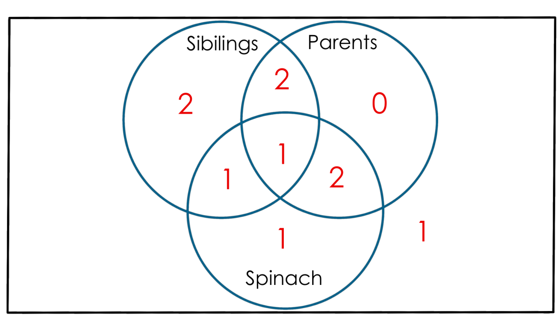
**Example** To simplify this example, I will only use ten people, instead of the 50 you need for the project, with the following questions:

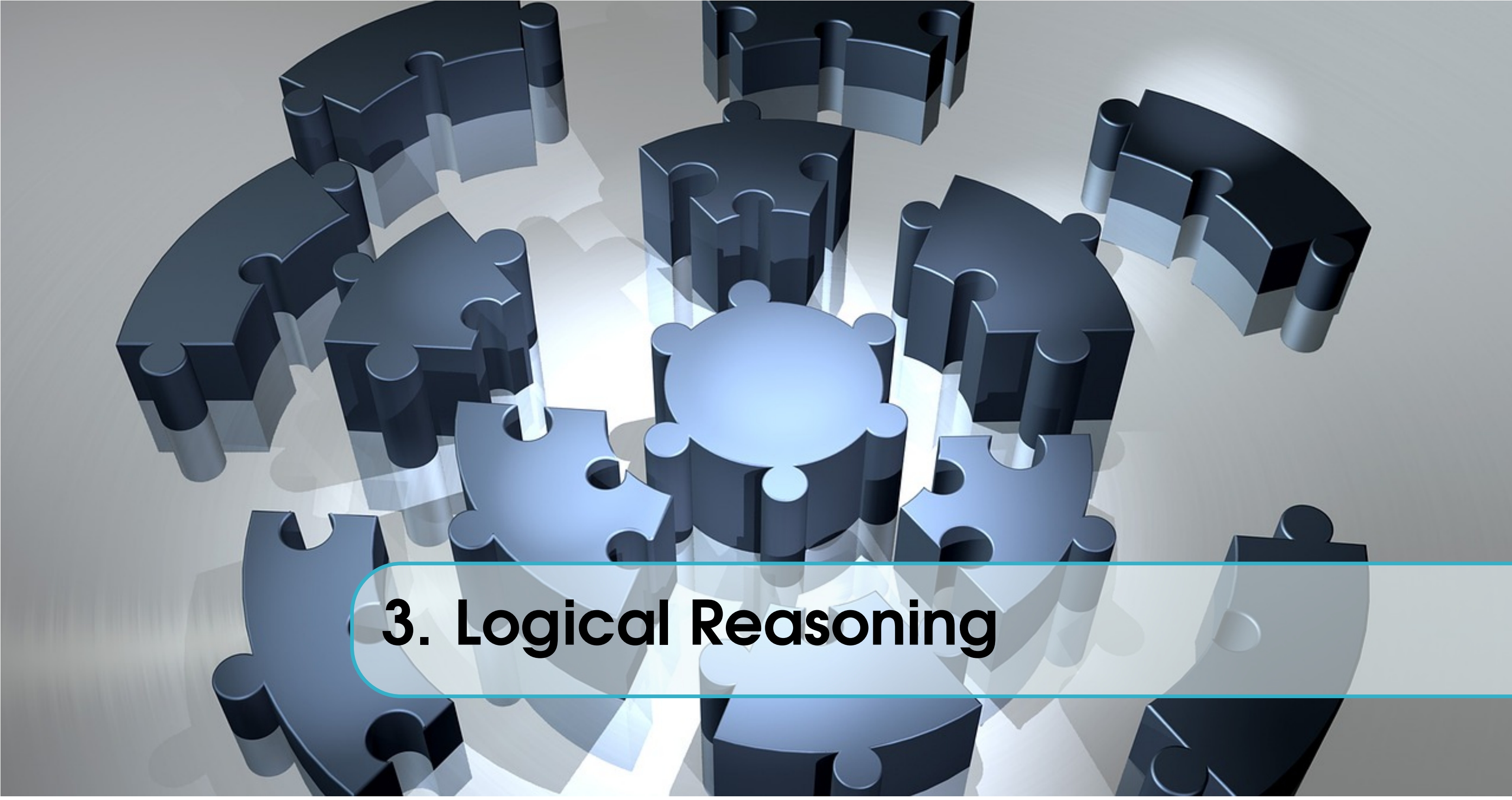
(1). Do you have any siblings?
(2). Do you live with your parents?
(3). Are you willing to eat spinach?

Here is the corresponding data-set:

| person | siblings? | live with parents? | eat spinach? |
|---|---|---|---|
| 1 | yes | yes | yes |
| 2 | no | no | yes |
| 3 | yes | no | no |
| 4 | no | yes | yes |
| 5 | no | yes | yes |
| 6 | yes | no | yes |
| 7 | no | no | no |
| 8 | yes | yes | no |
| 9 | yes | yes | no |
| 10 | yes | no | no |

Here is the corresponding Venn diagram:





# 3. Logical Reasoning

"The person who studies logic—law student, lawyer or judge—and who has become familiar with the principles of logical thinking is more likely to reason correctly than one who has not thought about the general concepts of reasoning." Ruggero J. Aldisert, Logic for Lawyers: A Guide to Clear Legal Thinking, at 23 (3d ed. 1997).

Have you ever wanted to find a loophole in a law? Have you ever wanted to get out of a contract? Have you ever wanted to prove something "beyond a shadow of a doubt"? The only way to accomplish these is to have an understanding of logic. The language of logic helps us to decide if our arguments are valid. It can help us determine whether the language of a contract or a law matches its intentions.

This chapter will define and use logic to understand basic types of sentences in the way a lawyer would think about them. Then we will take that further to applications of reasoning and surveys. The words we use will be simple (like AND or OR), but you may feel as if you are learning these words for the first time.

This chapter is inspired by the take from both the classic book in logic from Lewis Carroll **Symbolic Logic** (see [Car86]) as well the open-source textbook from Matthew J. Van Cleave **Introduction to Logic and Critical Thinking** (see [Cle16]).

We divide this chapter into two main sections. The first one is dedicated to the notion of propositional logic, observing the interaction with logical connectors and the representation of the proposition via the truth table. The second one will delve deeper into propositional logic via arguments, its presentation, and validation with Venn diagrams.

## 3.1 Propositions and Truth Values

**Definition 3.1.1 — Logic**

**Logic** is the science of the principles of reasoning.

We use logic in everyday life to help us solve problems. But to understand how to use logic correctly, we must first start with an understanding of what a proposition is.

**Definition 3.1.2 — Proposition**

A **proposition** is a complete sentence that states a claim. A proposition has a **truth value** which is either true or false.

■ **Example 3.1** The following sentences are propositions followed by their truth value.

| Proposition | Truth Value |
|---|---|
| Most dogs have four legs. | True |
| My dog is a cat. | False |
| $4+7=5$ | False |

Table 3.1: Examples of propositions and their truth values.

Some propositions will have a value of truth that depends on extra information. For example, the proposition "Today is Tuesday" will be either True or False, depending on the day of the week. ■

■ **Example 3.2** The following are not propositions:

| Statement | Why it is not a Proposition |
|---|---|
| Are you an English major? | This is a question. It does not claim anything. |
| 4+7 | This has no verb, so it is not a complete sentence. |
| The blue sky. | This has no verb, so it is not a complete sentence. |
| This sentence is false. | This sentence does not have a truth value. |

Table 3.2: Examples of sentences that are not propositions.

Notice each of these cases is not declaring something that can be associated with a value of truth. ■

We will sometimes name a proposition with a letter. For example:

$p$: My dog is a cat.
$q$: The sky is blue.

We can write the letter $p$ and know that it refers to the proposition, "My dog is a cat ."Similarly, $q$ refers to the proposition "The sky is blue."

When speaking, we can talk about opposites: short and tall, hot and cold, etc. However, in logic, it is vital to be more exact than this. We call this the **negation** of a proposition.

**Definition 3.1.3 — Negation.**
The **negation** of a proposition is another proposition that means precisely the opposite (always has the opposite truth value) of the original proposition.

- If a proposition $p$ is TRUE, then its negation, **not** $p$, is FALSE.
- If a proposition $p$ is FALSE then its negation, **not** $p$, is TRUE.

The notation for the negation of a proposition, $p$ is $\neg p$, for which we say "not $p$."

Since we call the negation of $p$ "not $p$," it follows that the way to negate a sentence is by using the word "not".

**Example 3.3** Consider the following proposition and negations.

- $p$: I am tall.
- $\neg p$: I am not tall

You can also negate a sentence by taking "not" out of the sentence. For example,

- $q$: I do not like candy corn.
- $\neg q$: I like candy corn.

■

Notice that the conclusion of this is that given a proposition $p$ then

$$\neg(\neg p) = p$$

or

| $p$ | $\neg p$ |
|---|---|
| T | F |
| F | T |

Table 3.3: Negation Truth Table.

Notice that any proposition is either True or False, so truth tables consider both cases.

■ **Example 3.4** The following are more examples of negations:

| **Proposition** | **Negation** |
|---|---|
| I am rich. | I am not rich. |
| I don't want to go to school. | I want to go to school. |
| $4+7=5$ | $4+7\neq 5$ |

Table 3.4: Propositions and Negations.

Notice that for each proposition in the table above, its negation has the opposite truth value. For example, if the statement, I am rich., is FALSE, then the statement, I am not rich, is TRUE, and vice versa. ■

Like many objects in mathematics, interactions between multiple elements always occur, and we can have compound sentences with one or more propositions.

### 3.1.1 Conjunction and Disjunction

When using the connectors AND and OR in logic, it is critical to define what they mean.

**Definition 3.1.4** Given two propositions $p$ and $q$ we define the **Conjunction** proposition as $p$ **AND** $q$. The compound proposition $p$ **AND** $q$ is true if and only if both $p$ is true and $q$ is true. $p$ AND $q$ can be written $p \wedge q$.

According to this definition, if any part of an AND statement is false, then the whole thing is false. To summarize this, we will use a truth table.

| $p$ | $q$ | $p \wedge q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | F |
| F | F | F |

Table 3.5: Conjunction Truth Table

■ **Example 3.5** Consider:

- $p$: Dollars is the currency in the US.
- $q$: Mordor is the capital of New Zealand.

The Conjunction proposition $p \wedge q$ would read, "Dollars is the currency in the US, and Mordor is the capital of New Zealand." This proposition is false as $p$ is True, but $q$ is False.
■

**Definition 3.1.5 — How to read a Truth table**
Notice how the table 3.5 is created.

- The first two columns show all of the possible combinations for $p$ and $q$. In the first row they are both true.
- In the second row $p$ is true and $q$ is false.
- In the third row $p$ is false and $q$ is true.
- And in the last row they are both false.
- Then the last column gives the result of the combinations of $p$ and $q$.

■ **Example 3.6** Determine if the following propositions are true or false.

(a). The earth is flat, and the sun is yellow.

(b) Dogs have four legs, and birds have two legs.

(c) Dogs have two legs, and birds have four legs.

(d) 4 is an even number, and 3 is an even number.

(e) Whales are mammals, and Dolphins are fish.

**Solution:**

(a). FALSE. p: "The earth is flat" is false. q: "The sun is yellow" is true. Since the only way, an AND statement can be true is if they are both true, $p \wedge q$ is FALSE.

(b) TRUE. Both parts of this sentence are true, so the whole sentence is true.

(c) FALSE. Both parts of this sentence are false. There is nothing true, so the whole sentence is false.

(d) FALSE. The first part of this sentence is true, but the second part is false. Both must be true for the whole thing to be true.

(e) FALSE. Both Whales and Dolphins are mammals, making the first sentence true and the second false, producing a false compound statement.

■

What if we had three propositions connected by AND? We use $r$ as our third proposition in the following truth table.

| $p$ | $q$ | $r$ | $p \wedge q \wedge r$ |
|---|---|---|---|
| T | T | T | T |
| T | T | F | F |
| T | F | T | F |
| T | F | F | F |
| F | T | T | F |
| F | T | F | F |
| F | F | T | F |
| F | F | F | F |

Table 3.6: Conjunction Truth Table for 3 Propositions

Notice that the only time $p \wedge q \wedge r$ is true is when all three are true.

**Definition 3.1.6** Given two propositions $p$ and $q$ we define the **Disjunction** proposition as $p$ **OR** $q$. The compound proposition $p$ **OR** $q$ is true if and only if either $p$ is true or $q$ is true or both are true.. $p$ OR $q$ can be written $p \vee q$.

According to this definition, as long as some part of an OR statement is true, the whole thing is true. To summarize this, we will use the following truth table.

| $p$ | $q$ | $p \vee q$ |
|---|---|---|
| T | T | T |
| T | F | T |
| F | T | T |
| F | F | F |

Table 3.7: Disjunction Truth Table for 2 Propositions

■ **Example 3.7** Consider:

- $p$: Dollars is the currency in the US.
- $q$: Mordor is the capital of New Zealand.

The Disjunction proposition $p \vee q$ would read:

- $p \vee q$: Dollars is the currency in the US or Mordor is the capital of New Zealand.

This proposition is True as $p$ is True and $q$ is False, but we only need one of them to be true for the compound proposition to be true. ■

Notice that the only time an OR statement is false is when all parts are false. We can extend this definition further to include a third proposition $r$.

| $p$ | $q$ | $r$ | $p \vee q \vee r$ |
|---|---|---|---|
| T | T | T | T |
| T | T | F | T |
| T | F | T | T |
| T | F | F | T |
| F | T | T | T |
| F | T | F | T |
| F | F | T | T |
| F | F | F | F |

Table 3.8: Disjuntion Truth Table for 3 Propositions

■ **Example 3.8** Determine if the following propositions are true or false.

(a) The earth is flat or the sun is yellow.

(b) Dogs have four legs or birds have two legs.

(c) Dogs have two legs or birds have four legs.

(d) 4 is an even number or 3 is an even number.

(e) Whales are mammals or Dolphins are fish.

**Solution:**

(a). TRUE. $p$: "The earth is flat" is false. $q$: "The sun is yellow" is true. Since one of them is true, the whole statement is true.

(b) TRUE. Both parts of this sentence are true, so the whole sentence is true.

(c) FALSE. Both parts of this sentence are false. There is nothing true, so the whole sentence is false.

(d) TRUE. The first part of this sentence is true, but the second part is false. Since some part of the statement is true, the whole statement is true.

(e) TRUE. Both Whales and Dolphins are mammals, making the first sentence true and the second false. Therefore producing a True compound statement.

■

### 3.1.2 Conditional Statements

In this subsection, we introduce another logical connector and evaluate the effect it produces in propositional statements.

**Definition 3.1.7** A **Conditional Statement** is any statement of the form if $p$, then $q$. In such a statement, $p$ is called the **hypothesis** of the statement and $q$ is called the **conclusion** of the statement.

■ **Example 3.9** For each of the following conditional statements, what is the hypothesis, and what is the conclusion?

- If I eat chocolate, then I am happy.
- If my dog is not happy, then I don't know what to do.

**Solution:**

- Premise: I eat chocolate. Conclusion: I am happy.
- Premise: My dog is not happy. Conclusion: I don't know what to do.

Notice that the words IF and THEN do not show up in any of these. ■

Analyzing the truth value of such a statement is similar to telling if someone is telling a lie or not.

■ **Example 3.10** Consider the following statement:

- If you find his fingerprints, then he is guilty.

Figure 3.1: Fingerprints on a surface

When is this statement a lie? Let's break this big proposition, if$p$ then $q$, into its smaller pieces $p$ and $q$. Therefore $p$: "You find his fingerprints." and $q$: "He is guilty."

- **Case 1: p and q are both true.** In this case, you find his fingerprints, and he is guilty. So the original statement, "If you find his fingerprints, then he is guilty," is true. In other words, you do not lie when you use it.

- **Case 2: p is true and q is false** In this case, you find his fingerprints, but he is not guilty. When you said, "If you find his fingerprints, then he is guilty," it was a lie and is therefore considered false.

- **Case 3: p is false and q is true** In this case, you don't find his fingerprints, and he is guilty. Does that mean you lied when you said, "If you find his fingerprints, then he is guilty"? No! Our conditional statement says nothing about what should happen if you don't find his fingerprints. It only says what happens if you do find them. So you have not lied, so the conditional statement is not false.

- **Case 4: p and q are both false** In this case, you don't find his fingerprints, and he is not guilty. Just like in case 3, our original conditional statement, "If you find his fingerprints, then he is guilty," says nothing about what should happen if you don't find his fingerprints. Hence, this has not made our conditional statement a lie.

■

The following truth table summarizes the truth values for If $p$, then $q$, which can be written $p \to q$.

| $p$ | $q$ | $p \to q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | T |
| F | F | T |

Table 3.9: Conditional Statement Truth Table

Notice that a conditional statement is false only when the premise is true and the conclusion is false.

■ **Example 3.11** Let:

- $p$: Dollars is the currency in the US.

- $q$: Mordor is the capital of New Zealand.

The conditional proposition $p \to q$ would read:

- $p \to q$: If Dollars is the currency in the US then Mordor is the capital of New Zealand.

This proposition is false as $p$ is true and $q$ is false. ■

### 3.1.3 Tautologies and Morgan Laws.

In this subsection, we introduce some properties of different operations between logical connectors and the notion of equivalence.

**Definition 3.1.8** Two propositions are **logically equivalent** if they always have the same truth values.

In other words, if two propositions have identical columns in a truth table, then they are considered to be logically equivalent. In the following example, we will use a truth table to show whether or not two propositions are logically equivalent.

■ **Example 3.12** Use a truth table to decide if $\neg(p \wedge q)$ is equivalent to $\neg p \vee \neg q$.

**Solution:**

To create this truth table, we will need columns for $p$ and $q$, but we will also need columns for any pieces along the way. For example, we will need to know the truth values for $\neg p$ and $\neg q$ before analyzing $\neg p \wedge \neg q$.

| $p$ | $q$ | $\neg p$ | $\neg q$ | | | |
|---|---|---|---|---|---|---|
| T | T | F | F | | | |
| T | F | F | T | | | |
| F | T | T | F | | | |
| F | F | T | T | | | |

Notice that every time $p$ is true, $\neg p$ is false and vice versa. The same holds for $q$ and $\neg q$. Now we can look at $\neg p \wedge \neg q$. To do this, look at the truth values of $\neg p$, and $\neg q$ in each row. Then decide what the truth value would be when connecting these two with AND.

| $p$ | $q$ | $\neg p$ | $\neg q$ | $\neg p \wedge \neg q$ | | |
|---|---|---|---|---|---|---|
| T | T | F | F | F | | |
| T | F | F | T | F | | |
| F | T | T | F | F | | |
| F | F | T | T | T | | |

The other proposition we need to evaluate is $\neg(p \vee q)$, but before we do, we need to have a column for $p \vee q$. Then we can negate that column.

| $p$ | $q$ | $\neg p$ | $\neg q$ | $\neg p \wedge \neg q$ | $p \vee q$ | $\neg(p \vee q)$ |
|---|---|---|---|---|---|---|
| T | T | F | F | F | T | F |
| T | F | F | T | F | T | F |
| F | T | T | F | F | T | F |
| F | F | T | T | T | F | T |

Now, we have to compare their columns to answer whether or not $\neg(p \vee q)$ is equivalent to $\neg p \wedge \neg q$. Looking at the table, both of these columns (the far right column and the third column from the right) are identical; this tells us that, yes, these are logically equivalent. ■

Notice that we also have that $\neg(p \wedge q)$ is equivalent to $\neg p \vee \neg q$ as the truth table below will show:

| $p$ | $q$ | $\neg p$ | $\neg q$ | $\neg p \vee \neg q$ | $p \wedge q$ | $\neg(p \wedge q)$ |
|---|---|---|---|---|---|---|
| T | T | F | F | F | T | F |
| T | F | F | T | T | F | T |
| F | T | T | F | T | F | T |
| F | F | T | T | T | F | T |

**Proposition 3.1.1 — De Morgan's law.** The propositions:

- $\neg(p \wedge q)$ and $\neg p \vee \neg q$
- $\neg(p \vee q)$ and $\neg p \wedge \neg q$

are logically equivalent.

Notice then negating that two things are happening simultaneously can be accomplished by showing one of them is not happening. Similarly, negating that either of two things is happening can be achieved by saying both are not happening.

**Definition 3.1.9** In propositional logic we have two special situations:

- A **Tautology** is a proposition for which all truth values are true.
- A **Contradiction** is a proposition for which all truth values are false.

If a proposition is a Tautology, its column would be all T's in its truth table. If a proposition is a Contradiction, then in its truth table, its column would be all F's.

Sometimes to check for the equivalency of two propositions or check for a Tautology is helpful to use another logical connector.

**Definition 3.1.10** The connector **Bi conditional** ($\Longleftrightarrow$) compares the values of truth of two propositions. Hence $p \Longleftrightarrow q$ is true if both $p$ and $q$ have the same truth values.

A summary of the behavior of the **Bi conditional** connector can be summarized as:

| $p$ | $q$ | $p \longleftrightarrow q$ |
|---|---|---|
| T | T | T |
| T | F | F |
| F | T | F |
| F | F | T |

Table 3.10: Bi-Conditional Statement Truth Table

■ **Example 3.13** Use the following truth table to show that $p \vee (p \to q)$ is a tautology.

| $p$ | $q$ | $p \to q$ | $p \vee (p \to q)$ |
|---|---|---|---|
| T | T | | |
| T | F | | |
| F | T | | |
| F | F | | |

**Solution:**

First, we need to use the information that we have in the values of the truth table for the conditional $p \to q$ then, we use the conjunction, logical operator, with the propositions $p$ and $p \to q$ to obtain the last column of truth values.

| $p$ | $q$ | $p \to q$ | $p \vee (p \to q)$ |
|---|---|---|---|
| T | T | T | T |
| T | F | F | T |
| F | T | T | T |
| F | F | T | T |

Since the final column is all T's, we know that $p \vee (p \to q)$ is a tautology. ■

■ **Example 3.14** Let:

- $p$: Dollars is the currency in the US.

- $q$: Dollars is not the currency in the US.

The proposition $p \longleftrightarrow \neg p$ would read:

- $p \longleftrightarrow q$: Dollars is the currency in the US if and only if Dollars is not the currency in the US.

This proposition is false as $p$ is true and $\neg p$ is false. Therefore not having the same truth values and proving not equivalent. ■

Of course, we can achieve equivalence between two propositions by showing the table for these propositions has the same truth values or by using a comparison via Bi-conditional that proves a tautology.

■ **Example 3.15** Use a truth table to show that:

- $\neg(p \to q)$
- $p \wedge \neg q$

are equivalent propositions.

**Solution:**

First, we must realize that to show that $p \wedge \neg q$ is the negation of $p \to q$, we are showing that they always have opposite truth values. If we fill in the table we get:

| $p$ | $q$ | $\neg q$ | $p \wedge \neg q$ | $p \to q$ |
|---|---|---|---|---|
| T | T | F | F | T |
| T | F | T | T | F |
| F | T | F | F | T |
| F | F | T | F | T |

Notice that the last two columns are opposites. Therefore they are negations of one-another. We can verify this further by extending the truth table and including the negation:

| $p$ | $q$ | $\neg q$ | $p \wedge \neg q$ | $p \to q$ | $\neg(p \to q)$ |
|---|---|---|---|---|---|
| T | T | F | F | T | F |
| T | F | T | T | F | T |
| F | T | F | F | T | F |
| F | F | T | F | T | F |

It is also common to use the logical connector **if and only if** or $\longleftrightarrow$ to compare two propositions' truth values in a single column. Equivalence proven by the **Tautology**.

| $p$ | $q$ | $\neg q$ | $p \wedge \neg q$ | $p \to q$ | $\neg(p \to q)$ | $(p \wedge \neg q) \longleftrightarrow \neg(p \to q)$ |
|---|---|---|---|---|---|---|
| T | T | F | F | T | F | T |
| T | F | T | T | F | T | T |
| F | T | F | F | T | F | T |
| F | F | T | F | T | F | T |

■

### 3.1.4 Converse, Inverse, and Contrapositive

When reading legal documents, people want to find loopholes (places where the logic of the record fails). Often these loopholes can be found when the document's author does not understand the different ways of rearranging conditional statements.

**Definition 3.1.11** For any conditional statement If $p$, then $q$, we can define the following:

- The **converse** of If $p$, then $q$ is: If $q$, then $p$.
- The **inverse** of If $p$, then $q$ is: If not $p$, then not $q$.
- The **contrapositive** of If $p$, then $q$ is: If not $q$, then not $p$.

A simple way of remembering this is that the converse switches $p$ and $q$; the inverse negates both $p$ and $q$, and the contrapositive does both.

■ **Example 3.16** Write the converse inverse and contrapositive of: If the moon is not full, then it is dark.

**Solution:**

- The **Converse:** If it is dark, then the moon is not full.
- The **Inverse:** If the moon is full, then it is not dark.
- The **contrapositive:** If it is not dark, then the moon is full.

Notice: All of the sentences still begin with the word IF. It is essential to realize that the parts of the sentence that change do not include the IF or THEN. ■

The big question is: Are any of these logically equivalent to the original sentence? To find out, we will use a truth table.

| $p$ | $q$ | $\neg p$ | $\neg q$ | $p \to q$ | $q \to p$ | $\neg p \to \neg q$ | $\neg q \to \neg p$ |
|---|---|---|---|---|---|---|---|
| T | T | F | F | T | T | T | T |
| T | F | F | T | F | T | T | F |
| F | T | T | F | T | F | F | T |
| F | F | T | T | T | T | T | T |

To decide what is equivalent to $p \to q$, we see if any of the other columns match its column exactly. We see that $\neg q \to \neg p$ has the same truth values for all entries; the contrapositive of an if...then... statement is logically equivalent to the original statement.

**A conditional statement and its contrapositive are logically equivalent to one another.** No other conditional statement is equivalent to the original. You may also notice that the converse and the inverse are equivalent to each other.

■ **Example 3.17** State if any of the following statements are equivalent to:

**If I don't sleep, then I can't stay awake.**

(a) If I sleep, then I can stay awake.

(b) If I can't stay awake, then I didn't sleep.

(c) If I can stay awake, then I sleep.

(d) If I can stay awake, then I didn't sleep.

**Solution:**

(a) No, because this is the inverse, not the contrapositive.

(b) No, because this is the converse, not the contrapositive.

(c) Yes, because this is the contrapositive.

(d) No, because in order for it to be the contrapositive, we have to negate BOTH the hypothesis and the conclusion.

■

■ **Example 3.18** For the following proposition, separate the hypothesis from the conclusion and write the converse, inverse, and contrapositive.

**"If it is raining, then the pavement is wet."**

**Solution:**

First, we can quickly separate the hypothesis and the conclusion.

**Hypothesis:** $p$ : it is raining

**Conclusion:** $q$ : the pavement is wet

Now

- The **Converse:** If the pavement is wet, then it is raining.
- The **Inverse:** If its not raining then the pavement is not wet.
- The **Contrapositive:** If the pavement is not wet, then it is not raining.

■

### 3.1.5 Exercises 3.1

**Problem 3.1** Decide if each of the following is a proposition. If not, state why.

(a) I am hot.

(b) The hot dog.

(c) Is the dog hot?

(d) My name is Frodo.

(e) I never tell the truth.

(f) To go to the store.

**Problem 3.2** Write the negation of each proposition.

(a) The boy is cold.

(b) That is not nice.

(c) California is in Europe.

(d) The moon is not in my pocket.

(e) Joseph goes to school.

**Problem 3.3** State whether each statement is TRUE or FALSE.

(a) $1+1=2$ or $1\times 1=1$.

(b) $1+1=2$ and $1\times 1=1$.

(c) Grass can be green and leaves must be purple.

(d) Grass can be green or leaves must be purple.

(e) Herbivores eat meat or carnivores eat meat.

(f) Herbivores eat meat and carnivores eat meat.

(g) The sun is cold and ice is hot.

**Problem 3.4** Write five new propositions and their negations.

**Problem 3.5** Fill in the truth table below and decide if $\neg p \vee q$ is equivalent to $\neg(p \vee q)$.

| $p$ | $q$ | $\neg p$ | $\neg p \vee q$ | $p \vee q$ | $\neg(p \vee q)$ |
|---|---|---|---|---|---|
| T | T | | | | |
| T | F | | | | |
| F | T | | | | |
| F | F | | | | |

**Problem 3.6** Fill in the truth table below and decide if $\neg p \vee \neg q$ is equivalent to $\neg(p \wedge q)$.

| $p$ | $q$ | $\neg p$ | $\neg q$ | $\neg p \vee \neg q$ | $p \wedge q$ | $\neg(p \wedge q)$ |
|---|---|---|---|---|---|---|
| T | T | | | | | |
| T | F | | | | | |
| F | T | | | | | |
| F | F | | | | | |

**Problem 3.7** For each of the following sentences, state the hypothesis and the conclusion.

(a) If the squirrel is clever, then the dog cannot catch it.

(b) If your brother is here, then you can see him.

(c) If you can't reach, then you are short.

(d) If the sky is not blue, then it is not sunny.

**Problem 3.8** For each sentence, write its converse, inverse, and contrapositive. Then state which of your new sentences is equivalent to the original sentence.

(a) If the squirrel is clever, then the dog cannot catch it.

(b) If your brother is here, then you can see him.

(c) If you can't reach, then you are short.

(d) If the sky is not blue, then it is not sunny.

**Problem 3.9** Use the truth table below to decide if $p \to q$ is logically equivalent to $\neg p \vee q$.

| $p$ | $q$ | $\neg p$ | $\neg p \vee q$ | $p \to q$ |
|---|---|---|---|---|
| T | T | | | |
| T | F | | | |
| F | T | | | |
| F | F | | | |

**Problem 3.10** Use the truth table below to decide if

$$[(p \to q) \to (q \to p)] \vee (p \vee q)$$

is a tautology.

| $p$ | $q$ | $p \to q$ | $q \to p$ | $(p \to q) \to (q \to p)$ | $p \vee q$ | $[(p \to q) \to (q \to p)] \vee (p \vee q)$ |
|---|---|---|---|---|---|---|
| T | T | | | | | |
| T | F | | | | | |
| F | T | | | | | |
| F | F | | | | | |

**Problem 3.11** Use the truth table below to decide if $(p \to q) \to r$ is logically equivalent to $p \to (q \to r)$.

| $p$ | $q$ | $r$ | $p \to q$ | $q \to r$ | $(p \to q) \to r$ | $p \to (q \to r)$ |
|---|---|---|---|---|---|---|
| T | T | T | | | | |
| T | T | F | | | | |
| T | F | T | | | | |
| T | F | F | | | | |
| F | T | T | | | | |
| F | T | F | | | | |
| F | F | T | | | | |
| F | F | F | | | | |

**Problem 3.12** Consider the propositions $P$ =*Paris is the capital of France* and $Q$= *Dollar is the currency of Japan*. State the truth value of

(a) $P \to Q$

(b) $P$ and $Q$

(c) $P$ or $Q$

**Problem 3.13** Write a definition for each of the following terms.

(a) and

(b) conditional statement

(c) contrapositive

(d) converse

(e) inverse

(f) logic

(g) logically equivalent

(h) negation

(i) not

(j) or

(k) premise

(l) proposition

(m) tautology

(n) truth value

**Problem 3.14** Which of the following sentences are propositions? If they are not propositions, say why.

(a) I don't want to grow up.

(b) Do you want to grow up?

(c) Grow up!

(d) To grow up.

(e) The brown cow.

(f) The cow is brown.

**Problem 3.15** Write the negation of each proposition.

(a) I am not in trouble.

(b) He is short.

(c) Life is complicated.

(d) Studying is not fun.

**Problem 3.16** Assume that "January is cold." is true and the proposition "December is hot" is false. State whether each of the following statements is true or false.

(a) January is cold and December is hot.

(b) January is cold and December is not hot.

(c) January is not cold and December is hot.

(d) January is not cold and December is not hot.

(e) January is cold or December is hot.

(f) January is cold or December is not hot.

(g) January is not cold or December is hot.

(h) January is not cold or December is not hot.

(i) If January is cold, then December is hot.

(j) If January is cold, then December is not hot.

(k) If January is not cold then December is hot.

(l) If January is not cold, then December is not hot.

**Problem 3.17** Write the converse, inverse, and contrapositive of each sentence. Then circle the sentence that is logically equivalent to the original sentence in each group.

(a) If I am a duck, then I say moo.

(b) If Bob is a human, then he is not from outer space.

(c) If the dog is not happy, then he will cry.

(d) If ferrets are not mammals, then they do not have fur.

(e) If I have a PlayStation 5, then I'm a gamer.

**Problem 3.18** Use the truth table below to decide if $\neg p \to (p \to q)$ is a tautology.

| $p$ | $q$ | $\neg p$ | $p \to q$ | $\neg p \to (p \to q)$ |
|---|---|---|---|---|
| T | T | | | |
| T | F | | | |
| F | T | | | |
| F | F | | | |

**Problem 3.19** Use the truth table below to decide if $(p \wedge q) \vee r$ is logically equivalent to $p \wedge (q \vee r)$.

| $p$ | $q$ | $r$ | $p \wedge q$ | $(p \wedge q) \vee r$ | $q \vee r$ | $p \wedge (q \vee r)$ |
|---|---|---|---|---|---|---|
| T | T | T | | | | |
| T | T | F | | | | |
| T | F | T | | | | |
| T | F | F | | | | |
| F | T | T | | | | |
| F | T | F | | | | |
| F | F | T | | | | |
| F | F | F | | | | |

**Problem 3.20** Use the truth table below to decide if $(p \vee q) \wedge r$ is logically equivalent to $p \vee (q \wedge r)$.

| $p$ | $q$ | $r$ | $p \vee q$ | $(p \vee q) \wedge r$ | $q \wedge r$ | $p \vee (q \wedge r)$ |
|---|---|---|---|---|---|---|
| T | T | T | | | | |
| T | T | F | | | | |
| T | F | T | | | | |
| T | F | F | | | | |
| F | T | T | | | | |
| F | T | F | | | | |
| F | F | T | | | | |
| F | F | F | | | | |

**Problem 3.21** Consider the propositions $P$ =*New York is the Capital of US* and $Q$= *Yen is the currency of Japan*. State the truth value of

(a) $P \longleftrightarrow Q$

(b) $P$ or $Q$

## 3.2 Logical Arguments and Venn Diagrams

In English, the word argument has many meanings, the most common of which makes a person think of fighting about something. In logic, we use a particular definition of argument aimed to create a chain of reasoning to support a claim.

**Definition 3.2.1** An **argument** is a set of statements, some of which (the **premises**) attempt to provide a reason for thinking that some other statement (the **conclusion**) is true.

**Example 3.19** A simple example of an argument would be:

- Premise: All people are happy.
- Premise: John is a person.
- Conclusion: Therefore, John is happy.

■

In some ways, this argument makes sense. If we know that all people are happy and that John is one of these people, he must be happy. But when you think about it, how can we say that all people are happy? In order to analyze this argument, we need to begin with a few definitions.

**Definition 3.2.2** An argument is **Valid** if the conclusion must follow from the premises.

This means that no matter what you think about the premises if they force the conclusion to be true, the argument is valid.

**Definition 3.2.3** An argument is **Sound** if it is valid and its premises are all true.

Looking back at the argument, John must be happy if all people are happy, and he is a person. So the argument is Valid. However, since the first premise is a lie (It is not true that everyone is happy), it is not a sound argument.

It turns out that you can use Venn Diagrams to help analyze statements' validity. While this seems simple to some, others will wonder how we know for sure if this argument is valid? It would be best if you were asking that question.

### 3.2.1 Using Venn Diagrams to Represent Sentences

A Venn Diagram is an illustration that uses circles to show the relationship among things or finite groups of things. Circles overlapping have commonality, while circles that do not overlap do not share traits. We often use Venn Diagrams to represent similarities and differing elements between ideas visually.

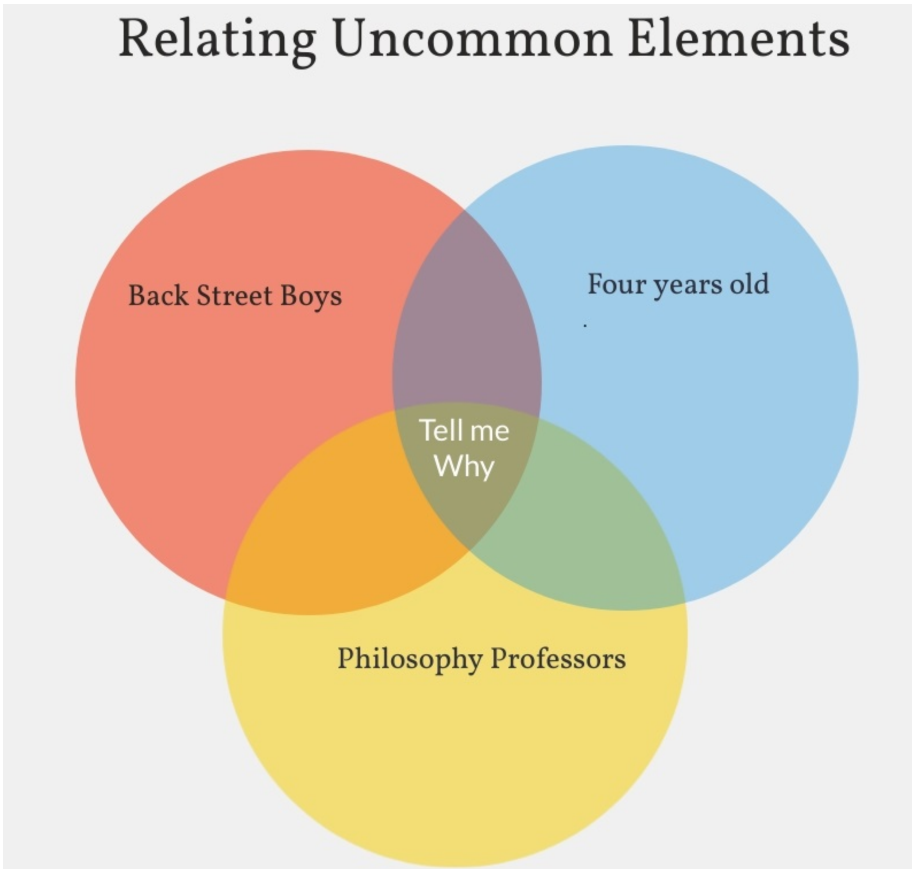



Figure 3.2: Expressing relationships through Venn diagrams

Analyzing arguments is helpful to be able to draw a picture of a premise. There are four types of sentences we will learn to draw:

- All Statements.
- No Statements.
- Some are.
- Some are not.

- **ALL statements**
  Let's start with the sentence: "All ice cream is cold." We need to know what our two circles in the Venn Diagram represent to draw this sentence. In this case, one of the sets will be ice cream, and the other set will be cold things. Since we are saying that all ice cream is cold, every kind of ice cream is in the collection of cold items. In other words, ice cream is a subset of cold things. So to draw this, we will put the circle of ice cream inside the circle of cold things:

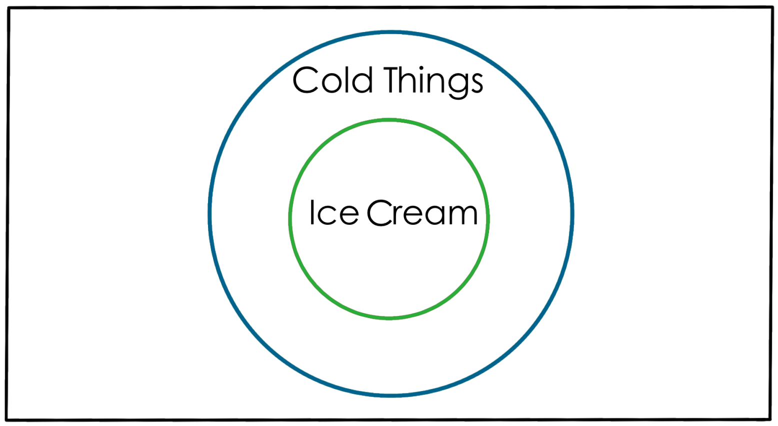



Figure 3.3: All ice cream is cold.

- **NO statements**
  Now, look at the sentence: "No ice cream is cold," this is telling you that there is nothing in common between these sets. In other words, their intersection is empty. So the best way to draw this would be for the two circles to be disjoint.

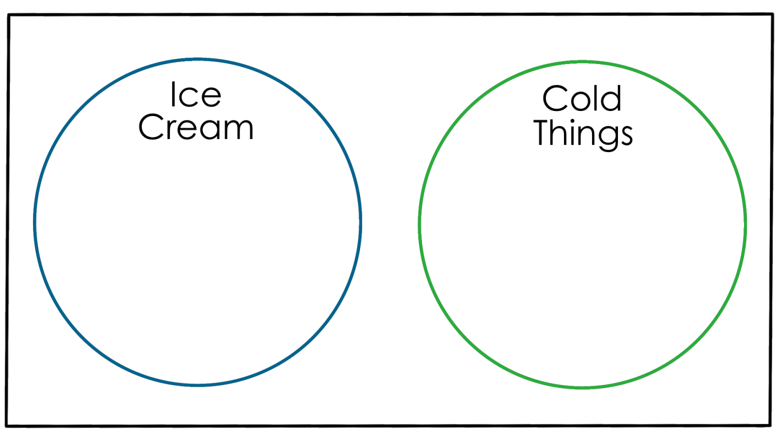



Figure 3.4: No ice cream is cold.

- **SOME ARE and SOME ARE NOT statements**
  The last two types of sentences are similar, so it is best if we look at them together.

  "Some ice cream is cold." and "Some ice cream is not cold."

  For SOME statements, we either know that our two sets do have something in their intersection, or there is something outside of their intersection. For this, we need overlapping circles. The difference is that in the sentence "Some ice cream is cold." we know there is something in the intersection. "Some ice cream is not cold" means that we know there is something in the part of the ice-cream circle that is not overlapping the cold circle. The only difference in these pictures is the placement of an x marking where the known object is.

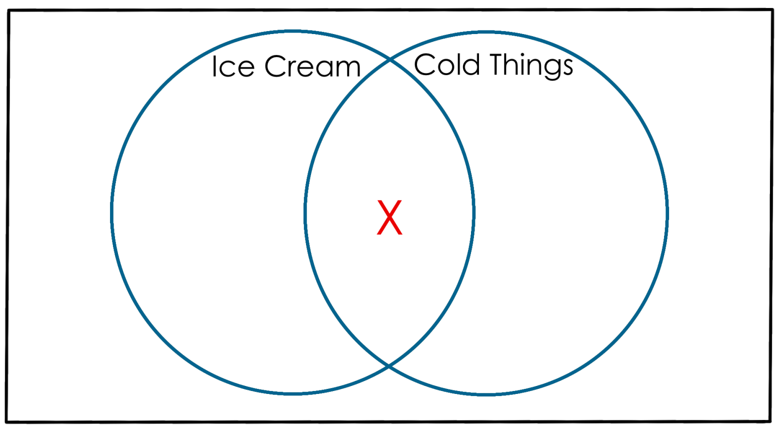



Figure 3.5: Some ice cream is cold.

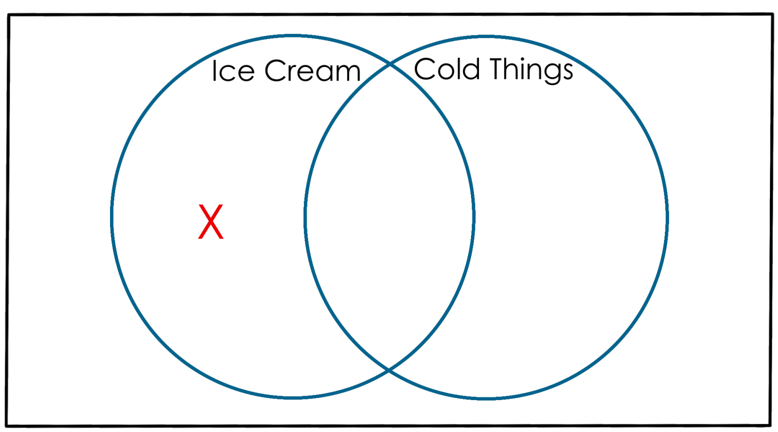



Figure 3.6: Some ice cream is not cold.

■ **Example 3.20** Draw a Venn diagram for each of the following sentences.

(a). All dogs are fast.

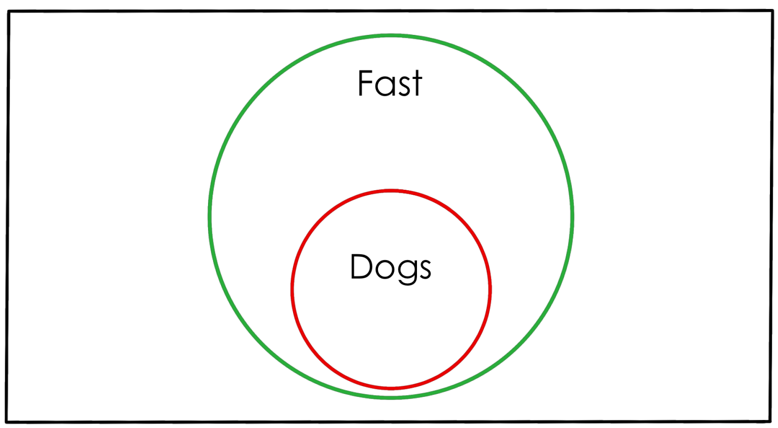
(b). No kids are mean.

(c). Some chocolate is dark.

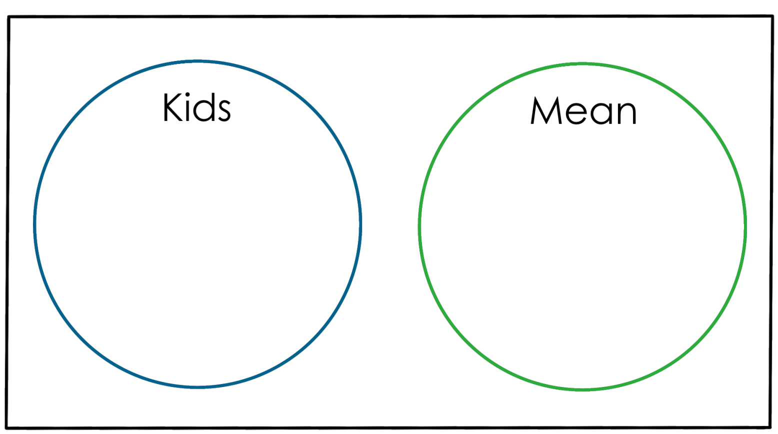
(d). Some chocolate is not dark.

**Solution:**

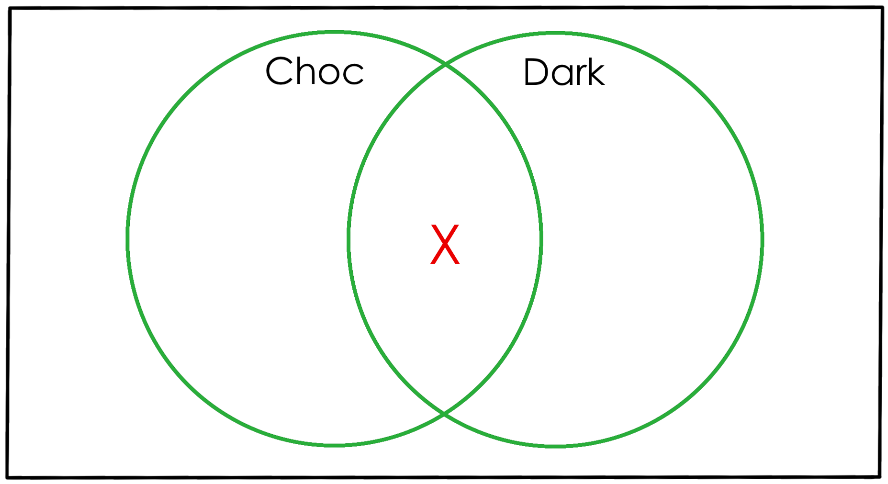
(a). Since all dogs are fast, all of the dogs are inside the circle of fast things.



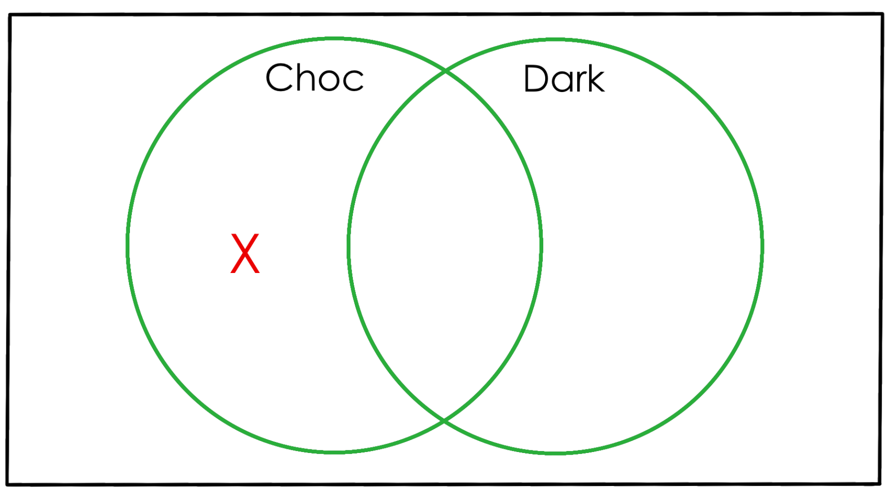

(b). Since none of the kids are in the mean circle, we will have separate circles.



- (c). This will be overlapping circles with an x showing that there is something in the intersection.




- (d). This will be the same picture as in part c, except the x shows that we know there is a chocolate outside of the dark circle.




■

Notice that this wording can be generalized to the statement that includes more than two elements and account for any type of relationships they may have.

### 3.2.2 Using Venn Diagrams to Analyze Arguments

Now that we know how to use Venn Diagrams to represent sentences, we can use them to analyze arguments. We will analyze statements using the following steps:

**How to Analyze the Validity of an Argument**

(1). Draw a Venn Diagram for the first premise; this should be the sentence with the word ALL, NO, or SOME.

(2). Draw an x inside the region of the diagram based on the second premise. If you could put the x in more than one section, draw it on the line between the sections.

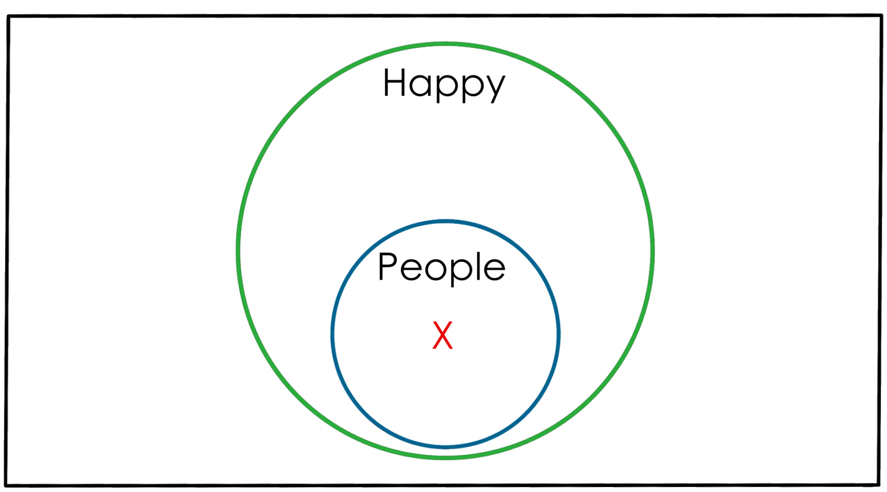
(3). Compare the conclusion to the picture. If they match, then the argument is Valid. If not, then it is not Valid.

Let's look back at the argument from the beginning of the section:

■ **Example 3.21** Use a Venn Diagram to show whether or not this argument is Valid or Sound.

- Premise: All people are happy.
- Premise: John is a person.
- Conclusion: Therefore, John is happy.

**Solution:**
We will draw the people's circle inside the happy circle based on the first premise. We will put an x, representing John, inside the people circle, based on the second premise.



Now we see if the picture shows that the conclusion must happen. In the picture, John is definitely inside the happy circle. The whole people circle is inside the happy circle. Therefore, this does show that John is happy. We can conclude that the argument is Valid.

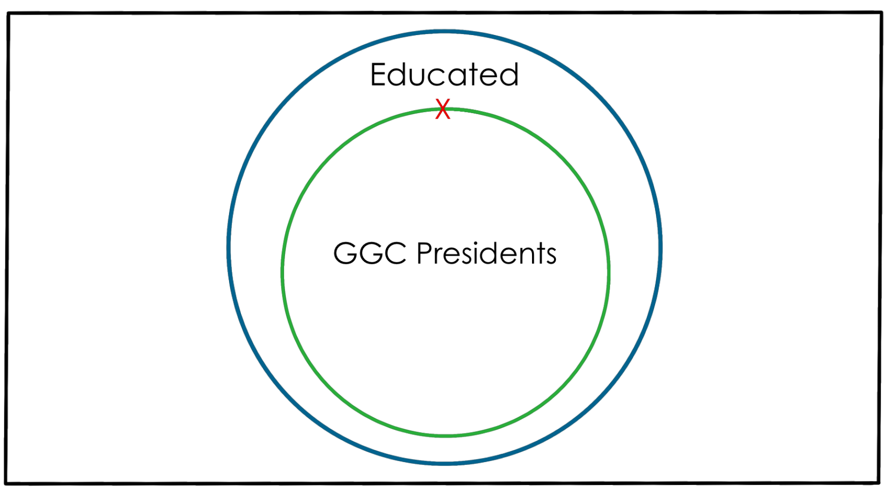
Since it is Valid, we need to check whether or not it is Sound. To be sound, both of the premises must be true. We already discussed that "All people are happy" is not a true statement. Therefore, we can conclude that the argument is Not Sound. ■

■ **Example 3.22** Use a Venn Diagram to show whether or not this argument is Valid or Sound.

- Premise: All GGC presidents have been educated.
- Premise: Dr. Joseph is educated.
- Conclusion: Therefore, Dr. Joseph is a GGC president.

**Solution**

We will draw the president's circle inside the educated circle based on the first premise. We will put an x, representing Dr. Joseph, inside the educated circle based on the second premise. There are two regions within the educated circle: the region inside the president's circle and the region outside it. Since the assumption does not say more than that the x is somewhere in the educated circle, we will put the x on the line separating the two regions.




Now we see if the picture shows that the conclusion must happen. The picture does not indicate that Dr. Joseph is a GGC president. Since she is on the line, it shows that she might be a GGC president or not be a GGC president. Since the picture does not match the conclusion exactly, we can conclude that the argument is Not Valid.

Since the argument is NOT Valid, the statement is automatically known as Not Sound.

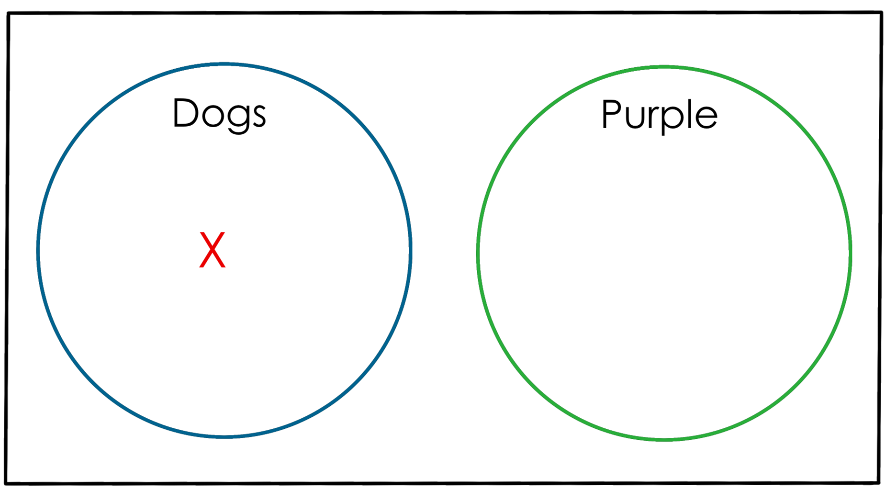
■

■ **Example 3.23** Use a Venn Diagram to show whether or not this argument is Valid or Sound.

- Premise: No dogs are purple.
- Premise: Fluffy is a dog.
- Conclusion: Therefore, Fluffy is not purple.

**Solution**

The dog circle and the purple circle should not touch based on the first premise. The second premise tells us to put Fluffy, represented as an x, inside the dog circle.
Now we see if the picture shows that the conclusion must happen. In the picture, the x is




outside the purple circle; this tells us that Fluffy is not purple. Therefore, the conclusion matches the picture, and we can conclude that the argument is Valid.

Since the argument is valid, we need to check the truth of the premises. It is true that no dogs are purple. We can assume Fluffy is a dog. So our hypotheses are true, and we can conclude that the argument is Sound. ■

### 3.2.3 Logical Arguments and Chain of Conditionals

Another way to analyze arguments is by using a chain of conditional statements. This method is excellent if you need to determine a conclusion to a set of premises.

**Definition 3.2.4** A **chain of conditionals** is a list of conditional statements where the conclusion of any argument is the hypothesis of the following statement.

■ **Example 3.24** An example of a chain of conditionals would be:

- If p, then q.
- If q, then r.
- If r, then s.
- If s, then t.
- If t, then u.

Notice that the first premise ends with q, and the second premise starts with q. The second premise ends with r, and the next one starts with r, and so on; this creates a chain:

$$p \to q \to r \to s \to t \to u$$

Because of this chain, the first proposition, $p$, must lead to the last one, $u$. So the Conclusion is: $p \to u$

■

■ **Example 3.25** What is the conclusion to the following argument:

- If I eat chocolate, then I am happy.
- If I am happy, then I work hard.
- If I work hard, then I get a raise.

**Solution:**

Since the premises are in an order that makes up a chain of conditionals, the conclusion will be: If I eat chocolate, then I get a raise. ■

Sometimes, the list of premises is not in a chain. In that case, you must rearrange them to make one.

■ **Example 3.26** What is the conclusion to the following argument:

- If I have not eaten, then I must cook breakfast.
- If I am not ready, then I will be late.
- If I woke after seven, then I have not eaten.
- If I must cook breakfast, then I am not ready.

**Solution**

To rearrange these, we must figure out where to start. Notice that two of the sentence pieces only occur once: I woke after seven, and I will be late. One of these must start the chain, and one must end it. Starting with I woke after seven:

- If I woke after seven, then I have not eaten.

  (The next sentence must start with I have not eaten.)

- If I have not eaten, then I must cook breakfast.

  (The next sentence starts with I must cook breakfast)

- If I must cook breakfast, then I am not ready.

  (The next sentence starts with I am not ready.)

- If I am not ready, then I will be late.

Since we started this chain with I woke after seven, and we ended it with I will be late, the conclusion is: If I woke after seven, then I will be late.

■

As you are doing problems like these, there are a lot of words that make them not as straightforward. One idea is to change each part of the sentence into a single letter, like $p$ or $q$. However, it is better to do this with letters that make sense for the situation.

In the following, we will redo the previous example with letters so that you can see another method of organization.

■ **Example 3.27** Redo the previous example using letters instead of words.

- If I woke after seven, then I have not eaten.
- If I have not eaten, then I must cook breakfast.
- If I must cook breakfast, then I am not ready.
- If I am not ready, then I will be late.

**Solution:**

Let's use letters that match the words:

- e = I have eaten (so I have not eaten would be not e)
- c= I must cook breakfast
- r= I am ready
- l= I will be late
- a= I woke after seven

Now we can rewrite our sentences more simply. By doing this, it is also easier to rearrange them into a chain of conditionals.

- If a, then not e
- If not e, then c
- If c, then not r
- If not r, then l

The Conclusion is If a, then l or "If I woke after seven, then I am late," which is the same answer we got in the previous example. In terms of symbols, this can be written as:

- $a \rightarrow e$
- $e \rightarrow c$
- $c \rightarrow r$
- $r \rightarrow l$

and the conclusion $a \rightarrow l$. ■

Now let's look at a small argument:

- If I like fruit, then I will eat an apple.
- If I don't like fruit, then I am hungry.

Since the word fruit comes up twice, we don't want to start with it. So somehow, we need to turn that first sentence around. We need to be careful because we know that the contrapositive of the conditional statement is the only correct way to rewrite it so that it has the same meaning.

So, "If I like fruit, then I will eat an apple." becomes "If I don't eat an apple, then I don't like fruit." Using this replacement for the first sentence, we now have:

- If I don't eat an apple, then I don't like fruit.
- If I don't like fruit, then I am hungry.

So the conclusion is: If I don't eat an apple, then I am hungry.

■ **Example 3.28** What is the conclusion to the following argument:

- If I hide the lamp, then my mom can't find it.
- If I am not in trouble, then my mom can find it.
- If I didn't hide the lamp, then I didn't break the lamp.

**Solution:**

We need to start by finding words that only occur once. Even though "hide the lamp" and "don't hide the lamp" are slightly different (in fact, opposites), hiding the lamp is mentioned twice in some way, so we will not start with that phrase. The two parts that occur only once are "I am not in trouble" and "I didn't break the lamp ."So we need to start with one of those. I choose to begin with, "I didn't break the lamp." So I will start with the contrapositive of the third statement.

- If I break the lamp, then I hide the lamp.

  (now we need to start with "I hide the lamp")

- If I hide the lamp, then my mom can't find it.

(Since we need to start with "If my mom can't find it," we will use the contrapositive of the second sentence).

- If my mom can't find it, then I am in trouble.

Now that we have a chain, we can conclude, "If I break the lamp, then I am in trouble." ■

One question you may have about the last example is what would have happened if I started with "I am not in trouble"? If I had done that, my conclusion would have been the contrapositive of the conclusion we found in the example:

- If I am not in trouble, then I didn't break the lamp.

Since the two statements are logically equivalent, they are both correct answers.

**Changing ALL or NO Statements into Conditional Statements.**


Figure 3.7: Transformation of statements will be key in checking if arguments are valid and sound (Photo by Ross Findom).

A critical fact about conditional statements that we overlooked in previous sections is that we can rewrite ALL and NO statements as IF...THEN... statements.

■ **Example 3.29** Consider the proposition:

**All mice are blind.**

What does this tell you about something called a mouse? It says:

**If that is a mouse, then it is blind.**

Any statement of the form "All p are q" is equivalent to "If p then q." ■

It takes a bit of thought to complete the statement transformation for NO statements.

■ **Example 3.30** What does it mean for us to say:

**No balloon is orange.**

Since the word **no** is involved, we will need to negate something, but which part? Does this sentence tell us anything about objects that are not balloons? No! So we will not negate the first part. Instead, think about what this is saying about balloons. It is saying that:

**If it is a balloon, then it is not orange.**

Any statement of the form "No p are q" is equivalent to "If p then not q." ■

Let's see some more examples of transforming conditional statements.

■ **Example 3.31** Write each of the following sentences as a conditional statement.

(a). All tall boys can reach the top.

(b). All candy is not good for you.

(c). No babies cry.

(d). No elephants are not large.

**Solution:**

(a). If you are a tall boy, then you can reach the top.

(b). If it is candy, then it is not good for you.

(c). If he is a baby, then he won't cry.

(d). If that is an elephant, then it is large.

■

Now that we can change ALL and NO statements to conditional statements, we can come up with a conclusion involving them.

■ **Example 3.32** Change each statement to a conditional statement. Then use the conditional statements to form a chain and find the conclusion.

- All who don't have dry eyes are crying.
- All people with friends are happy.
- No happy people cry.

**Solution**

We start by changing each sentence to an equivalent conditional statement.

- If a person does not have dry eyes, then he is crying.
- If a person has friends, then he is happy.
- If a person is happy, then he will not cry.

Now we look to see what is only mentioned once. That would be dry eyes and having friends. We can start with either to get one of two correct answers. I am randomly choosing to start with the first sentence.

- If a person does not have dry eyes, then he is crying.

  (Now we must start with "crying," so we will need to use the contrapositive of the third sentence.)

- If he is crying, then he is not happy.

  (Now we must start with "not happy," so we must use the contrapositive of the second sentence.)

- If he is not happy, then he has no friends.

Now that we have a chain of conditionals, we can find the conclusion:

- If a person does not have dry eyes, then he has no friends.

  or No people without dry eyes have friends.

If we had started with "having friends," we would have ended up with the contrapositive of our conclusion:

- If he has friends, then he has dry eyes, or All people with friends have dry eyes.

■

### 3.2.4 Exercises 3.2

**Problem 3.22** Write a definition for each of the following terms.

- (a) and
- (b) argument
- (c) chain of conditionals
- (d) conclusion of an argument
- (e) conclusion of a conditional statement
- (f) conditional statement
- (g) sound
- (h) valid

**Problem 3.23** Draw a Venn diagram representing each sentence below. If the sentence starts with SOME, make sure to include an x in the correct region.

(a) All girls are interesting.

(b) All moms are cool.

(c) No fish can fly.

(d) Some canteloupe is ripe.

(e) Some buffalo are big.

(f) No man is an island.

(g) Some people don't like squash.

(h) Some cats are not black.

**Problem 3.24** Using a Venn diagram, show whether or not the argument is valid. Then state whether or not it is Sound.

Premise: All fish can swim.

Premise: The elephant cannot swim.

Conclusion: So, the elephant is not a fish.

**Problem 3.25** Using a Venn diagram, show whether or not the argument is valid. Then state whether or not it is Sound.

Premise: All music is fun.

Premise: Hip-hop is fun.

Conclusion: So, Hip-Hop is music.

**Problem 3.26** Using a Venn diagram, show whether or not the argument is valid. Then state whether or not it is Sound.

Premise: All dogs are large.

Premise: Pluto is a dog.

Conclusion: So, Pluto is large.

**Problem 3.27** Using a Venn diagram, show whether or not the argument is valid. Then state whether or not it is Sound.

Premise: No monkeys are mammals.

Premise: You are a monkey.

Conclusion: So, you are a not a mammal.

**Problem 3.28** Using a Venn diagram, show if the argument is Valid or Sound.

Premise: No students will fail.

Premise: You are a student.

Conclusion: So, you will not fail.

**Problem 3.29** Using a Venn diagram, show if the argument is Valid or Sound.

Premise: Some dogs bite.

Premise: Snoopy is a dog.

Conclusion: So, Snoopy bites.

**Problem 3.30** Write each of the following sentences as a conditional statement.

(a). All miles are long.

(b). No leaves are not beautiful.

(c). All people that have cars have transportation.

(d). No astronauts are grounded.

**Problem 3.31** Write each of the following as sentences that begin with All or No.

(a). If a dog goes to sleep, then it will snore.

(b). If a cat does not sleep, then it is mad.

(c). If the government is wise, then we will not start wars.

(d). If the world is not clean, then we cannot breathe.

**Problem 3.32** Find the conclusion to the following argument.

If the world is not round, then you will fall off the edge.

If you don't float through space, then you do not fall off the edge.

If you float in space, then you are weightless.

**Problem 3.33** Find the conclusion to the following argument.

If the mouse ate the cheese, then the cat didn't eat the mouse.

If the dog did not stop the cat, then the cat ate the mouse.

If the dog is on a leash, then the dog did not stop the cat.

**Problem 3.34** Find the conclusion to the following argument.

If money is tight, then we are not wealthy.

If we are not wealthy, then we must work hard.

If money is not tight, then the economy is good.

**Problem 3.35** Find the conclusion to the following argument.

No human is an alien.

All non-humans are not from Earth.

No one who isn't from Earth has seen a dog.

**Problem 3.36** Find the conclusion to the following argument.

No damaged cars run.

All cars that are not broken are good enough.

All good enough cars are cars I would buy.

**Problem 3.37** Find the conclusion to the following argument.

No unpaid clowns get laughs.

All talented clowns are funny.

All funny clowns get laughs.

**Problem 3.38** Rewrite each sentence as a conditional statement.

(a). All mice are tiny.

(b). No boys want to sit.

(c). No friends want to not see each other.

(d). All candidates are not honest.

(e) All Video-games are violent.

**Problem 3.39** Draw a Venn diagram to represent the sentence. If it is a "Some" sentence, remember to include an x.

(a). All cantaloupe are ripe.

(b). No cantaloupe is ripe.

(c). Some cantaloupe are ripe.

(d). Some cantaloupe are not ripe.

(e). Some mammals are aquatic.

**Problem 3.40** Rewrite the sentences as a chain of conditionals. Then write the conclusion.

Premise: If monkeys fly, then the sky is blue.

Premise: If monkeys don't fly, then it is cold outside.

Premise: If it is not winter, then it is not cold outside.

**Problem 3.41** Draw a Venn diagram based on the premises, and then tell if the argument is Valid or Sound.

Premise: All music is peaceful.

Premise: Hard rock is music.

Conclusion: Hard rock is peaceful.

**Problem 3.42** Draw a Venn diagram based on the premises, and then tell if the argument is Valid or Sound.

Premise:No chocolate is sour.

Premise:The candy is not sour.

Conclusion: The candy is chocolate.

**Problem 3.43** Draw a Venn diagram based on the premises, and then tell if the argument is Valid or Sound.

Premise: All dogs are mammals.

Premise: A poodle is a dog.

Conclusion: A poodle is a mammal.

**Problem 3.44** Rewrite the sentences as a chain of conditionals. Then write the conclusion.

No green things are food.

All money is green

All things that are not food make me sad.

# 3.3 Project

## 3.3.1 Project: Lewis Carroll Logic Poems

### Introduction

Lewis Carroll, the author of Through the Looking Glass (renamed Alice in Wonderland), once said:

"Once master the machinery of Symbolic Logic, you have a mental occupation always at hand, of absorbing interest, and one that will be of real use to you in any subject you may take up. It will give you clearness of thought, the ability to see your way through a puzzle, the habit of arranging your ideas in an orderly and get-at-able form, and more valuable than all, the power to detect fallacies and to tear to pieces the flimsy illogical arguments, which you will so continually encounter in books, in newspapers, in speeches, and even in sermons, and which so easily delude those who have never taken the trouble to master this fascinating Art." (Carroll, 1896)

Carroll was not only a great writer, but he was also an accomplished mathematician. Much of what he wrote seemed nonsensical and silly, but if you analyze it and create chains of conditionals, you will find there is a hidden conclusion to everything he wrote.

The purpose of this project is for you to use your knowledge of symbolic logic to find the hidden meaning in Lewis Carroll's Logic poems.

### Instructions

First, pick two poems from Poem Set A and two verses from Poem Set B. Then, for each poem, do the following:

(a). Change each statement to If... then... form.

(b). Label the two parts of each sentence with letters to write them in the shorter form: If p, then q.

(c). Line them up so that they make a chain. Don't forget you may need to use the contrapositive.

(d). Write the conclusion in letters.

(e). Write the conclusion in words using if and then.

(f). Write the conclusion in a "more poetic" way.

### Example

Each of your chosen poems should be done in this way. Notice that some words have the same meaning, and some have opposite meanings. For example, you can change short to not tall. While we have learned that short and not tall are not logically equivalent, we will say they are for use in these poems. (That's the beauty of poetry).

All tall folks can reach the top.
All short folks are slow.
No early folks are slow.

(a). Change each sentence to If...then...form.

If you are tall, then you can reach the top.
If you are short, then you are slow.
If you are early, then you are fast

(b). Label the two parts of each sentence with letters to write them in If p then q form.

t = you are tall

r = you can reach the top

not t = you are short (since short is the same as not tall)

s = you are slow

e = you are early

not s= you are fast (since fast and not slow are the same.)

If t then r.
If not t then s
If e then not s

(c). Line them up so that they make a chain.

In this case, we can either start with r or e since they are the only letters that you only see once. Since e is already at the beginning of a line, I choose to start with e.

If e then not s
If not s then t
If t then r

(d).Write the conclusion in letters.

If e then r

(e). Write the conclusion in words.

If you are early, then you can reach the top.

(f). Write the conclusion in a more "poetic" way.

Look back at the wording of the original poem for an idea of what this should look like.

All early folks can reach the top.

**Poem Set A:**
Consider the following verses and choose two.

All puddings are nice.
This dish is a pudding.
No nice things are wholesome.

Everyone who is sane can do logic.
No lunatics are fit to serve on a jury.
None of your sons can do logic.

All unripe fruit is unwholesome.
All these apples are wholesome.
No fruit grown in the shade is ripe.

All hummingbirds are richly colored.
No large birds live on honey.
Birds that do not live on honey are dull in color

No potatoes of mine that are new have been boiled.
All my potatoes in this dish are fit to eat.
No unboiled potatoes of mine are fit to eat.

My saucepans are the only things I have that are made of tin.
I find all your presents very useful.
None of my saucepans are of the slightest use.

**Poem Set B:**

Consider the following verses and choose one.

All writers, who understand human nature, are clever.
No one is a true poet unless he can stir the hearts of men.
Shakespeare wrote "Hamlet".
No writer, who does not understand human nature, can stir the hearts of men.
None but a true poet could have written "Hamlet."

No boys under 12 are admitted to this school as boarders.
All the industrious boys have red hair.
None of the day boys learn Greek.
None but those under 12 are idle.

The only books in this library that I do not recommend for reading are unhealthy in tone.
The bound books are all well written.
All the romances are healthy in tone.
I do not recommend you to read any of the unbound books.

Promise breakers are untrustworthy.
Wine drinkers are very communicative.
A man who keeps his promises is honest.
No teetotalers are pawnbrokers.
One can always trust a very communicative person.

All my sons are slim.
No child of mine is healthy and takes no exercise.
All gluttons, who are children of mine, are fat.
No daughter of mine takes any exercise.

The only articles of food that my doctor allows me are such as are not very rich.
Nothing that agrees with me is unsuitable for supper.
A wedding cake is always very rich.
My doctor allows me all articles of food that are suitable for supper.

I despise anything that cannot be used as a bridge.
Everything worth writing an ode to would be a welcome gift to me.
A rainbow will not bear the weight of a wheelbarrow.
Whatever can be used as a bridge will bear the weight of a wheelbarrow.
I would not take, as a gift, a thing that I despise.

# 4. Probability

What comes to mind when you think about "probability"? You may think about "odds" or the "chance of something occurring ."Probability is the mathematical study of randomness, chance, and uncertainty. Probability involves asking questions about uncertain events: What is the likelihood it will rain today? What is the chance your favorite basketball team will win the championship game? How likely is it you will win the lottery?

The historical development of probability is based on gambling and games of chance. The applications of probability have since expanded to numerous fields. In modern times, we may use the probability within the healthcare field to assess the risks of specific treatments or predict the chance of getting a particular disease. Insurance companies use probability to determine the likelihood of being in an automobile accident based on geographical region. Stockbrokers use probability to predict the potential rate of return on their investments. Although there are many ways we use chance, we classify probability into three main types:

- Theoretical
- Experimental
- Subjective

This chapter has been inspired by everyday situations used for probability. Some exercises have been extracted from the Quantitative Reasoning Workbook (see [BD20]).

**Definition 4.0.1 Theoretical probability** involves the mathematical calculation of probabilities. We do not have to physically toss a coin to predict the probability of the coin landing on heads. In this chapter, much of our focus is on theoretical probability.

**Definition 4.0.2 Experimental probability** (or empirical probability) differs from theoretical probability because we conduct experiments under controlled conditions to determine chances. If we toss a coin 50 times and record the number of times it lands on heads and the number of times it lands on tails, we can use the results to calculate the experimental probability.

■ **Example 4.1** Let's imagine the experiment of throwing a coin 20 and then 50 times.


Figure 4.1: Theoretical chances in coin-tossing may differ from the actual experimental result.

The theoretical probability of heads and tails is $P(H) = P(T) = 0.50$ or that each case has half the chance. Now consider the following experimental values obtained by tossing a coin 20 times:

| Outcome | Frequency |
|---|---|
| Head | 13 |
| Tails | 7 |
| Total | 20 |

Table 4.1: The coin toss experiment 50 times

Notice that the frequency for heads is $P(H) = \frac{13}{20} = 0.65$ or a 65% while the frequency for tails is $P(T) = \frac{7}{20} = 0.35$ or 35%, of course our coin is well balanced. Now let's increase the number of tosses we use for the experiment to 50; this produces the following results.

| Outcome | Frequency |
|---|---|
| Head | 26 |
| Tails | 24 |
| Total | 50 |

Table 4.2: The coin toss experiment 50 times

Now the frequency for heads is $P(H) = \frac{26}{20} = 0.52$ or a 52% while the frequency for tails is $P(T) = \frac{24}{20} = 0.48$ or 48%. Notice that although the values are much closer to the theoretical value, they are not the same. ■

**Definition 4.0.3** The third type of probability is **subjective probability**. This type of probability is based on our intuitions or experiences. Subjective probability does not involve calculations. Instead, it is based on our judgment of the likelihood of an event occurring. For example, if you regularly arrive early for appointments, you might speculate there is an 85% chance you will arrive early for your next dentist appointment. In this chapter, we will focus only on theoretical and experimental probability.

■ **Example 4.2** An example of subjective probability would be the statement:


Figure 4.2: Probabilities that have no experimental or theoretical basis are subjective.

Other examples would be:

- "As your professor, I'm 100% sure you will like your grade in this class."
- "I feel there is a 50-50 chance of rain today."

■

## 4.1 Theoretical Probability

We need some basic terminology as we begin our exploration of probability:

**Definition 4.1.1** Let's assume that we have a situation that can be observed:

- The result of an observation or experiment is called an **outcome**.
- A **sample space** is the set of all possible outcomes.
- An **event** is any particular outcome or group of outcomes.

### 4.1.1 Expressing Probabilities

We denote the probability of an event, $A$, as $P(A)$. If all of our outcomes are equally likely, we calculate the probability of event $A$ as

$$P(A) = \frac{\text{number of ways an event can occur}}{\text{total number of possible outcomes}}$$

■ **Example 4.3 — Coin tossing.** A simple example is the tossing of a fair coin. One possible outcome is the coin lands on heads, H. The second possible outcome is the coin lands on tails, T. For one coin toss, the sample space is

$$S = \{H, T\}$$

. If we are interested in the event of the coin landing on heads, the probability is

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

■

■ **Example 4.4 — 6-sided die.** If we roll a 6-sided die, what is the probability of the event of landing on a 4? There are six possible outcomes which give us a sample space of

$$S = \{1, 2, 3, 4, 5, 6\}$$

The die has only one 4, so the probability of landing on a 4 is

$$P(4) = \frac{1}{6}$$

What if we are interested in the event of getting an outcome greater than 4 with a roll of the same die? There are two outcomes greater than 4, so the probability is

$$P(\text{outcome greater than } 4) = \frac{2}{6} = \frac{1}{3}$$

■

■ **Example 4.5** If we choose, at random, a family with three children, what is the probability the family has two boys and one girl? The sample space for three children is

$$S = \{\text{BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG}\}.$$

There are eight possible outcomes, but the event of two boys and one girl occurs three times: BBG, BGB, GBB. Therefore, the probability is

$$P(2 \text{ boys, } 1 \text{ girl}) = \frac{3}{8}$$

■

■ **Example 4.6 — Role-playing dice.** Polyhedral dice go beyond the usual six sides. We have 4-side, 8-side, 10-side, 12-side, and 20-side dice.


Figure 4.3: Polyhedral dice

Thanks to that, we can consider the following probabilities.

(a) The chance of getting a 7 in 10-side die.

(b) The chance of getting a sum of 6 when tossing two 4-side dice.

**Solution:**

(a) Since 7 is one outcome of the possible 10, $\{1, 2, .., 10\}$, then $P(7) = 1/10$.

(b) The set of outcomes to obtain a sum of 6 are given by the pair of rolls

$$\{(3,3),(4,2),(2,4)\}$$

Since we have 4-side dice, we have 16 possible sums in total, therefore

$$P(\text{Sum of } 6) = \frac{3}{16}$$

.

■

■ **Example 4.7 — Playing Cards.** Playing cards are often used in probability problems. A standard deck of playing cards contains 52 cards. The deck has four **suits** of 13 cards:

- Clubs
- Spades
- Hearts
- Diamonds

Clubs and spades are black, and hearts and diamonds are red. Each card suit contains one ace, nine cards numbered 2 - 10, a jack, a queen, and a king (see figure 4.4). Note most standard decks of cards include two jokers, but we will not have those here.

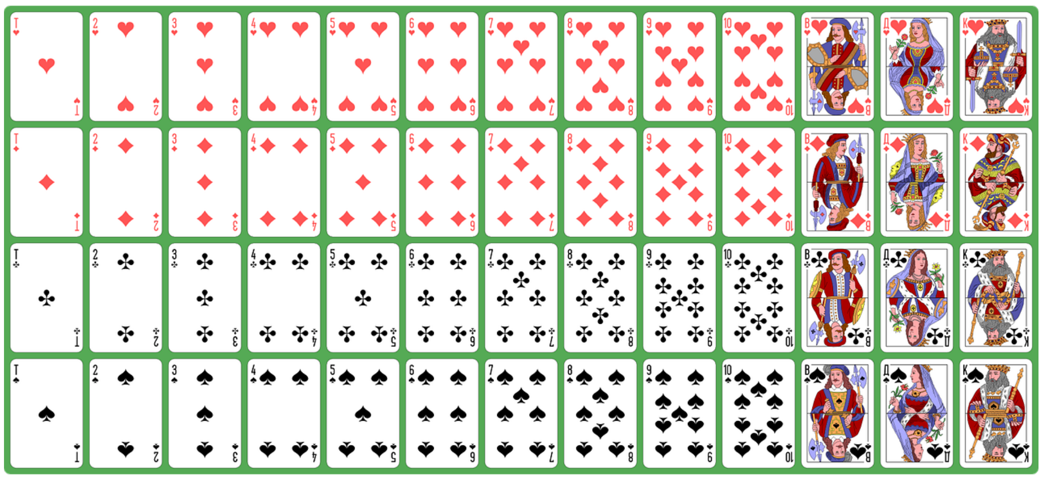


Figure 4.4: A deck of cards

Thanks to that, we can consider the following probabilities.

(a) What is the probability of randomly drawing one card from a deck and getting a jack?

(b) What is the probability of randomly drawing one card from a deck and getting a red card?

**Solution:**

(a) There are 52 cards and 4 of them are jacks, so the probability is

$$P(\text{Jack}) = \frac{4}{52} = \frac{1}{13}$$

(b) Since we have 26 red cards out of 52

$$P(\text{Red}) = \frac{26}{52} = \frac{1}{2}$$

■

### 4.1.2 Probability Values

Notice the answers in our examples are fractions. We express the probability of an event using a number between 0 and 1. We can use simplified fractions, decimals, or percents. An event with a probability of 0 is considered *impossible*. An event with a probability of 1 is considered *certain*. The majority of probabilities lie somewhere between 0 and 1. A higher probability means an event is more likely to occur than one with a lower probability. The probability of any event must be $0 \leq P(A) \leq 1$ .

### 4.1.3 Complement of an Event

What if we wanted to know the probability that an event does not happen? If we see the likelihood of an event, we can use that information to help us.

■ **Example 4.8** We found in a previous example that the probability of rolling a 4 with a six-sided die is

$$P(4) = \frac{1}{6}$$

.

What is the probability of not getting a 4? From our sample space for a 6-side die,

$$\{1,2,3,4,5,6\}$$

We can see five outcomes are not a four; therefore,

$$P(not\ a\ 4) = \frac{5}{6}$$

We can also express this answer as a decimal, $0.8\bar{3}$, or as a percent, 83.3%.

■

**Definition 4.1.2** Finding the probability of an event not occurring involves finding the **complement** of the event. The complement of any event, $A$, is the probability event "A does not happen ."We denote the complement of the event $A$ as $A'$. We calculate the complement of A as

$$P(A') = 1 - P(A)$$

■ **Example 4.9 — Cards.** What is the chance of not getting an ace if we are dealt a card from a deck?

**Solution:**
Since we have an ace for each of the four suits, the probability of getting an ace is

$$P(A) = \frac{4}{52}$$

Therefore, not obtaining an ace can be found using the complement formula.

$$P(A') = 1 - P(A) = 1 - \frac{4}{52} = \frac{48}{52}$$

■

### 4.1.4 Calculating Odds

You may have heard the term *odds* used in place of the terms *chance* or *probability*. While odds are a measure of the likelihood of a particular event, they are calculated differently than the methods we have used so far. Odds are expressed as ratios and used in gambling and statistics applications.

Figure 4.5: Odds are famously featured in horse races.

**Definition 4.1.3** **Odds** are ratios used to calculate the chance of an event occurring by comparing the number of ways an event can occur to the number of ways the event cannot.

The odds *in favor* of an event, $A$, is the ratio of the probability of $A$ to its complement.

$$\frac{P(A)}{P(A')} = \frac{P(A)}{1-P(A)}$$

The odds *against* an event, $A$, is the ratio of the complement of $A$ to the probability of $A$.

$$\frac{P(A')}{P(A)} = \frac{1-P(A)}{P(A)}$$

Now let's see some examples of the use of odds.

■ **Example 4.10** The morning weather report shows there is an 80% chance of rain today. That means there is a 20% chance it will not rain. The odds of it raining today are

$$\frac{P(\text{rain})}{P(\text{no rain})} = \frac{0.80}{1-0.80} = \frac{0.80}{0.20} = \frac{4}{1}$$

This means we are four times more likely to get rain today than not. ■

■ **Example 4.11** You buy a lottery ticket that has a 1 in 10 chance of being a winning ticket. The odds against having a winning ticket is

$$\frac{P(\text{not winning})}{P(\text{winning})} = \frac{10-1}{10} = \frac{9}{10}$$

This means you have a 90% chance of not having a winning ticket. ■

■ **Example 4.12** A horse is considered to have 9-2 odds.

(a) Explain what this means with odds formula.

(b) If you place a bet of $2 dollars in the horse, what would be your payoff?

**Solution:**

(a) Since the odds are 9-2 this means that from a total of 11 possible races, the horse could win 9, hence the probability of the horse winning would be:

$$P(A) = 9/11 = 81.82\%$$

Notice if we use this in the formula for odds in favor, we obtain the ratio.

$$\frac{P(A)}{P(A')} = \frac{P(A)}{1-P(A)} = 9/2 = 4.50$$

This is also the proportion that will define your earnings.

(b) A horse that wins at 9-2 odds will return $4.50 for every $1.00 wagered. If you had placed the minimum bet of $2 on that horse to win, your payoff would be:

$$(4.50x\$2) + \$2$$

The last two dollars from the original value of the bet, for a total of $11.

■

### 4.1.5 Experimental Probability

Figure 4.6: In experimental probability we need to know how often certain situations occur.

**Definition 4.1.4** **Experimental** (or *empirical*) probability differs from theoretical probability because we are calculating the probability of future events based on collected observations or experiments.

Experimental probability is helpful when we do not have the information needed to calculate a theoretical probability, such as predicting the chance of rain or the likelihood a particular basketball player will make a free throw. When we use the outcomes of past observations or experiments, we can estimate the probability of future events.

We calculate experimental probability as follows.

$$P(A) = \frac{\text{observed number of times A occurs}}{\text{total number of observed occurrences}}$$

**Definition 4.1.5** This calculation is also called the **relative frequency** of the event.

■ **Example 4.13** Travel survey 100 people were asked how they got to work:

- 37 used a car
- 12 rode a bicycle
- 42 took public transport
- 9 walked

**Solution:**

The experimental probability of picking someone going to work using one of the modes of transportation would be

- Car $= \frac{37}{100}$
- Public Transport $= \frac{42}{100}$
- Bicycle $= \frac{12}{100}$
- Walking $= \frac{9}{100}$

■

■ **Example 4.14** Insurance companies often use empirical probability to determine insurance rates for drivers in a particular state, such as Georgia. They might estimate the probability of a licensed driver being in an automobile accident by calculating the probability using existing data:

Figure 4.7: In 2015, there were 382,043 automobile accidents in Georgia (`https://gdot.numetric.net/crash-data`) among 6,906,191 licensed drivers (`https://www.fhwa.dot.gov/policyinformation/statistics/2015/dl1c.cfm`)

This data can be used to estimate the probability of being in an accident in Georgia, using the following conventions for the events.

- $L =$ Licensed driver being in an automobile accident
- $N =$ Number of accidents involving licensed drivers in Georgia
- $G =$ Number of licensed drivers in Georgia

$$P(L) = \frac{N}{G} = \frac{382,043}{6,906,191} = 0.055 = 5.5\%$$

would give the probability of licensed driver being in an automobile accident. ■

■ **Example 4.15** The theoretical probability of flipping a coin and getting tails is 0.5. What if we flipped a coin 100 times? Would we always call 50 heads and 50 tails? We would most likely get different numbers of heads and tails. Imagine we toss a coin 100 times and get 42 heads and 58 tails. This means the experimental probability of getting tails is

$$P(\text{tails}) = \frac{58}{100} = 0.58$$

The relative frequency for getting tails in this example is 58%. ■

The relative frequency in the above example is different from what we would expect based on theoretical probability. How do we reconcile this difference? If we flip the coin many times, the relative frequency value will get closer to the theoretical value; this is known as the **law of large numbers**.


Figure 4.8: A Simulation of coin tossing with 200 coins. The chart shows how probability eventually settles in the theoretical probability for obtaining heads. `https://digitalfirst.bfwpub.com/stats_applet/stats_applet_10_prob.html`

**Definition 4.1.6** The **law of large numbers** states that as the number of repetitions of an experiment increases, the relative frequency obtained in the experiment gets closer to the theoretical probability.

### 4.1.6 Exercises 4.1

For each of the following problems, express each of your probabilities as both fraction and a decimal rounded to three places.

**Problem 4.1** Write the sample space for rolling one die. Find the probability of the event of rolling one dice and getting a number less than 4.

**Problem 4.2** If we randomly pick a month of the year:

(a) List the sample space, $S$.

(b) Find the probability that we pick a month that starts with a "J".

**Problem 4.3** You have a bag containing 15 apples, five are green, and ten are red. If you pick an apple randomly, what is the probability that it will be red?

**Problem 4.4** Find the probability of randomly drawing a red ace from a well-shuffled deck of cards.

**Problem 4.5** Find the probability of not drawing a red ace from a well-shuffled deck of cards.

**Problem 4.6** Fill in the following table for the sample space of all the possible sums from rolling two dice. Then find the probability of getting a sum of 7.

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | ____ | ____ | ____ | ____ | ____ | ____ |
| 2 | ____ | ____ | ____ | ____ | ____ | ____ |
| 3 | ____ | ____ | ____ | ____ | ____ | ____ |
| 4 | ____ | ____ | ____ | ____ | ____ | ____ |
| 5 | ____ | ____ | ____ | ____ | ____ | ____ |
| 6 | ____ | ____ | ____ | ____ | ____ | ____ |

**Problem 4.7** If a pair of fair dice is rolled once, what are the odds in favor of rolling a sum of 5?

**Problem 4.8** If a pair of fair dice is rolled once, what are the odds against rolling a sum of 7?

**Problem 4.9** If we roll two 6-sided dice, find the following:

1. The probability of getting a sum less than 7?
2. The probability of rolling a double?
3. The probability of not rolling a double?

**Problem 4.10** The table shown below gives the body weight for 156 office employees in a certain company.

| Body Weight (lbs) | Males | Females | Totals |
|---|---|---|---|
| 140-159 | 2 | 28 | 30 |
| 160-179 | 7 | 23 | 30 |
| 180-199 | 18 | 15 | 33 |
| 200-219 | 19 | 7 | 26 |
| 220-239 | 17 | 3 | 20 |
| 240-259 | 15 | 2 | 17 |

If we randomly pick an office employee, what is the probability that it will be:

1. A female.
2. A person weighing between 200 and 219 lbs.
3. A male weighing between 160 and 199 lbs.

**Problem 4.11** Bags of dog food are labeled 1 through 50. What is the probability that the dog will choose a bag that is not numbered with a multiple of 11?

**Problem 4.12** If you draw a single card from a standard deck, what are the odds in favor of drawing a face card?

**Problem 4.13** A basketball team has a 45% probability of winning. Find the odds against them winning.

**Problem 4.14** The table below gives the age of 120 adult residents in a particular neighborhood together with information if they own or rent their home or apartment.

| Age | Rent | Own | Totals |
|---|---|---|---|
| 20-29 | 22 | 6 | 28 |
| 30-39 | 10 | 26 | 36 |
| 40-49 | 8 | 32 | 40 |
| 50-59 | 4 | 12 | 16 |
| Totals | 44 | 76 | 120 |

If a random adult resident of this neighborhood is chosen, find the probability it will be:

1. A person who rents his/her home/apartment.
2. A person between 20 and 29 years old.
3. A homeowner between 30 and 49 years old.

**Problem 4.15** You have a bag of fruit containing 12 apples, five oranges, and eight lemons. If you drop the bag and one of the fruits falls out, what are the odds against it being an orange?

**Problem 4.16** Consider the event of rolling an even number when rolling one standard die. List all the outcomes for this event.

**Problem 4.17** Consider the event of rolling a five and a 3 when rolling two standard dice. List all the outcomes for this event.

**Problem 4.18** Determine the probabilities of the following outcomes and events:

1. Tossing two coins and getting either one or two heads.
2. Randomly selecting a three-child family with exactly two girls.

**Problem 4.19** Suppose we toss a pair of standard twelve-side dice.

1. What is the probability that we get a 7?
2. What is the probability that we get any sum but 7?

**Problem 4.20** If we roll two dice 100 times and get a sum of 4 on 20 rolls, what is the relative frequency of getting a sum of 4? Express your answer as a percent. Is the relative frequency less than or greater than the theoretical probability of getting a sum of 4 with a roll of two dice?

**Problem 4.21** An independent theater sold 30 candy bars out of a total of 78 snacks last month. Based on this data, find the probability the next snack sold is a candy bar.

**Problem 4.22** A teacher makes the following list of the grades she gave to her 25 students on an essay:

$$A, C, C, B, C, D, C, C, F, D, C, C, C, B, B, A, B, D, B, A, A, B, F, C, B$$

Complete the following table and answer the questions that follow.

| Grade | Relative Frequency |
|---|---|
| A | |
| B | |
| C | |
| D | |
| F | |
| Total | |

1. What is the relative frequency of A?
2. What is the relative frequency of B?
3. What is the relative frequency of F?

**Problem 4.23** Investigate the concept of the **gambler fallacy** and explain.

1. How does the law of the large numbers prove this concept?
2. Find the probability of winning the Roulette game in a casino.
3. Apply the law of large numbers and explain who wins after many attempts.

## 4.2 Multistage Probability and Counting

We are often interested in calculating the likelihood of two or more events occurring. However, we must consider if the events could occur at the same time or if the occurrence of one event will affect the occurrence of another event. In this section, we will learn to calculate different types of multistage probabilities.

### 4.2.1 Independent and Dependent Events

Suppose we flip a coin and roll a die and want to know the probability of getting a tail on the coin and a two on the die. We could list all possible outcomes:

$$\{(H,1),(H,2),(H,3),(H,4),(H,5),(H,6),(T,1),(T,2),(T,3),(T,4),(T,5),(T,6)\}$$

and find

$$P(\text{Tails and } 2) = P(T,2) = \frac{1}{12}$$

Tossing a coin does not affect the outcome of rolling a die. These are called **independent events**.

**Definition 4.2.1** Two or more events are considered **independent** if the occurrence of one event does not affect the probability of the second (or third or fourth, etc.) event.

If two events, $A$ and $B$, are independent, then their combined probability is

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

In our previous example, "using a coin and a die," we could multiply the individual probabilities instead of listing all the possible outcomes. Check it yourself!

■ **Example 4.16** You have two bags with four marbles. Each bag contains one blue, one red, one black, and one yellow marble. What is the probability of choosing a red marble from one bag and a yellow marble from the other?

$$P(\text{red and yellow}) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$$

■

Sometimes we need to compute the probability of an event given that another event has occurred. Suppose you want to know the likelihood of drawing two aces from a standard deck of cards without putting the first one back into the deck. Although you begin with 52 cards, once you draw a card from the deck, there are now 51 cards remaining. The probability of drawing a second ace has changed by drawing an ace and not replacing it in the deck.

**Definition 4.2.2** Two or more events are considered **dependent** if the occurrence of one event changes the probability of the second (or third or fourth, etc.) event.

If two events, $A$ and $B$, are dependent, then their combined probability is

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

In our previous example of drawing two aces from a deck of cards without replacing the first card, once we draw an ace, three aces remain out of 51 cards, so the probability is

$$P(\text{ace and ace}) = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{13} \cdot \frac{1}{17} = \frac{1}{221}$$

■ **Example 4.17** You have a bag of 10 marbles. 4 are green, three are red, and three are blue. Find the probability of drawing three blue marbles in a row without returning any of the marbles to the bag.

$$P(\text{blue and blue and blue}) = \frac{3}{10} \cdot \frac{2}{9} \cdot \frac{1}{8} = \frac{6}{720} = \frac{1}{120}$$

■

■ **Example 4.18** Given a deck of cards, we are dealt two cards without returning them to the deck. Find the probability that we obtain two kings.

**Solution:**

Notice that initially we have 4 kings in the 52 cards deck, hence $P(\text{King}) = \frac{4}{52}$. Now for the second card, we have only three kings remaining and 51 cards; hence the probability of two kings is equivalent to the probability of getting a king, given you got a king from the first event.

$$P(\text{King and King}) = P(\text{King}) \cdot P(\text{King given King}) = \frac{4}{52} \cdot \frac{3}{51}$$

■

### 4.2.2 Inclusive and Exclusive Events

When considering multiple events and their probabilities, the natural question is to ask if they are both happening simultaneously or if one automatically prevents the other from occurring.

Figure 4.9: Notice the event of being 9:15 am is an exclusive event with the event of being another time, but it can be inclusive with the events of being a rainy or humid day.

**Definition 4.2.3** When two events, $A$ and $B$, can occur, but not simultaneously, we say they are **mutually exclusive**.

In the case of mutually exclusive events, we find the probability of event $A$ *or* event $B$. For example, if we roll one die, we cannot get two and a four simultaneously. These events are mutually exclusive. We can, however, find the probability of getting a two or a 4.

If two events, $A$ and $B$, are mutually exclusive, the probability of either event occurring is the sum of the individual probabilities.

$$P(A \text{ } or \text{ } B) = P(A) + P(B)$$

In our previous example of rolling one die and getting a 2 or a 4, the probability is

$$P(2 \text{ } or \text{ } 4) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

■ **Example 4.19** What is the probability of drawing either a red six or a black eight on one draw from a standard deck of cards? These two events are mutually exclusive because we cannot draw a red six and a black eight at the same time, so the probability is

$$P(\mathit{red}\ 6\ \mathit{or}\ \mathit{black}\ 8) = \frac{2}{52} + \frac{2}{52} = \frac{4}{52} = \frac{1}{13}$$

■

What about multiple events that share some characteristic or outcome? We must be careful when calculating these probabilities, so we do not double count the shared outcomes. For example, what if we wanted to know the likelihood of getting a diamond or a jack from a well-shuffled deck of 52 cards? The event of getting a diamond and the occurrence of getting a jack share the outcome of getting a jack of diamonds. If we calculated the individual probabilities, we would double count the one card that is a jack of diamonds.

**Definition 4.2.4** When two events, $A$ and $B$, have at least one common outcome, we say they are **non-mutually exclusive**.

If two events, $A$ and $B$, are non-mutually exclusive, the probability of either event occurring is the sum of the individual probabilities minus the shared outcomes.

$$P(A\ \mathit{or}\ B) = P(A) + P(B) - P(A\ \mathit{and}\ B)$$

In our previous example of getting a diamond or a jack from a well-shuffled deck of cards is

$$P(\text{diamond or jack}) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

■ **Example 4.20** If you roll two dice, what is the probability that we roll a total that is even or more than 8? The sample space for getting an even sum when rolling two dice is

$$\{2,4,6,8,10,12\}$$

and the sample space for getting a sum more than 8 is

$$\{9,10,11,12\}$$

These outcomes share the sums of 10 and 12, so our probability is

$$P(\mathit{even\ sum\ or\ sum\ greater\ than}\ 8) = \frac{6}{36} + \frac{4}{36} - \frac{2}{36} = \frac{12}{36} = \frac{1}{3}$$

■

The ability to count is fundamental to learning mathematics. It is something we know early in life. But what if we want to count how many ways we can arrange different items of interest?

What if we wanted to know how many ways we can organize a team of 3 friends from a group of 8? What if we wanted to know how many meals we could make if we had four different appetizers and three main courses? These questions require counting methods beyond just using numbers and adding up the different arrangements.

### 4.2.3 Fundamental Counting Principle

The basic idea says that when we are counting, we can add and multiply to summarize the quantities; taking this into account, let's start with some basic counting examples.

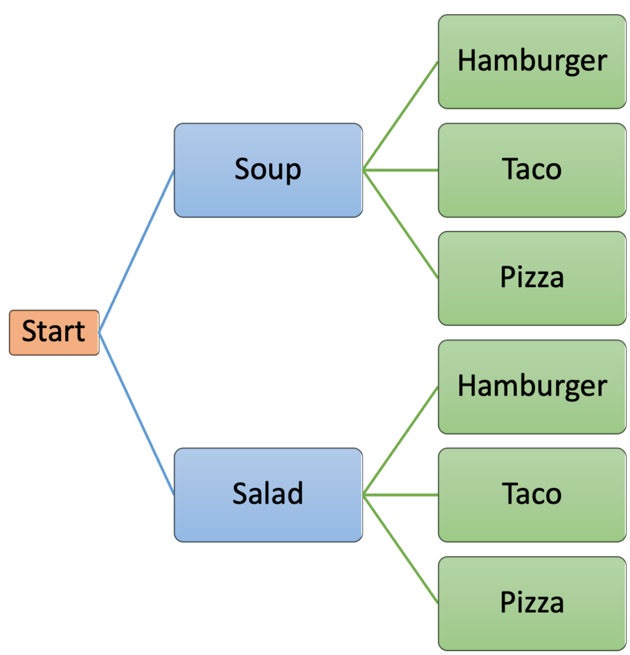
■ **Example 4.21** Suppose at a particular restaurant you have two choices for an appetizer (soup or salad) and three for the main course (hamburger, taco, or pizza). If you are allowed to choose exactly one item from each category for your meal, how many different meal options do you have?

**Method One**: One method for solving this problem is to list all of the possible meal options:

- Soup and hamburger
- Soup and taco
- Soup and pizza
- Salad and hamburger
- Salad and taco
- Salad and pizza

This method is not always efficient because we might skip a particular pair or list a pair twice. Assuming that we did this systematically and that we neither missed any possibilities nor listed any chance more than once, the answer would be 6. Thus you could go to the restaurant six nights in a row and have a different meal each night.

**Method 2**: Yet another method is to use a tree diagram.



**Method 3**: Another method we could use is to list all possible combinations in a table.

| | Hamburger | Taco | Pizza |
|---|---|---|---|
| **Soup** | Soup and hamburger | Soup and taco | Soup and pizza |
| **Salad** | Salad and hamburger | Salad and taco | Salad and pizza |

We can see how to count possibilities using lists, tables, and tree diagrams. These methods will continue to be helpful in some instances but imagine a game where you have two decks of cards (with 52 cards in each deck) and select one card from each deck. Would you want to draw a table or tree diagram to determine the number of outcomes of this game?

■

Notice in the above example that we could make six combinations no matter which method we used. You may also notice we could see the answer by simply multiplying two times 3.

We can generalize this technique as the sometimes called **Multiplication Principle** or also the **Fundamental Counting Principle**.

**Definition 4.2.5** The **Fundamental Counting Principle** states if we are asked to choose one item from each of two (or more) separate categories where there are $m$ items in the first category and $n$ items in the second category, then the total number of available choices is $m \cdot n$

■ **Example 4.22** We can extend our previous example to include dessert and use the Fundamental Counting Principle to find the total number of choices. Suppose at a particular restaurant you have three choices for an appetizer (soup, salad, or breadsticks), five choices for the main course (hamburger, sandwich, quiche, taco, or pasta), and two options for dessert (pie or cake).

If you can choose exactly one item from each category for your meal, how many different meal options do you have? There are

$$3 \cdot 5 \cdot 2 = 30 \text{ possible options}$$

■

■ **Example 4.23** Consider choosing a car from either **Honda, Mazda, Subaru, or Kia** brands. We also can pick a color from **gray, blue, red, orange, black, or yellow** and with two different models **sedan or hatchback**. By applying the counting principle, let's calculate how many possible combinations you can choose from.

$$4 \cdot 6 \cdot 2 = 48 \text{ possible options}$$

■

### 4.2.4 Permutations and Combinations

What if we want to count all possible arrangements using the letters from the word MATH? This problem involves counting arrangements rather than the ways we can combine one item from two or more different categories (such as our restaurant example).

We could create a list of all possible arrangements, which would be tedious and inefficient. Moreover, the question only asks for *how many* arrangements we can make.

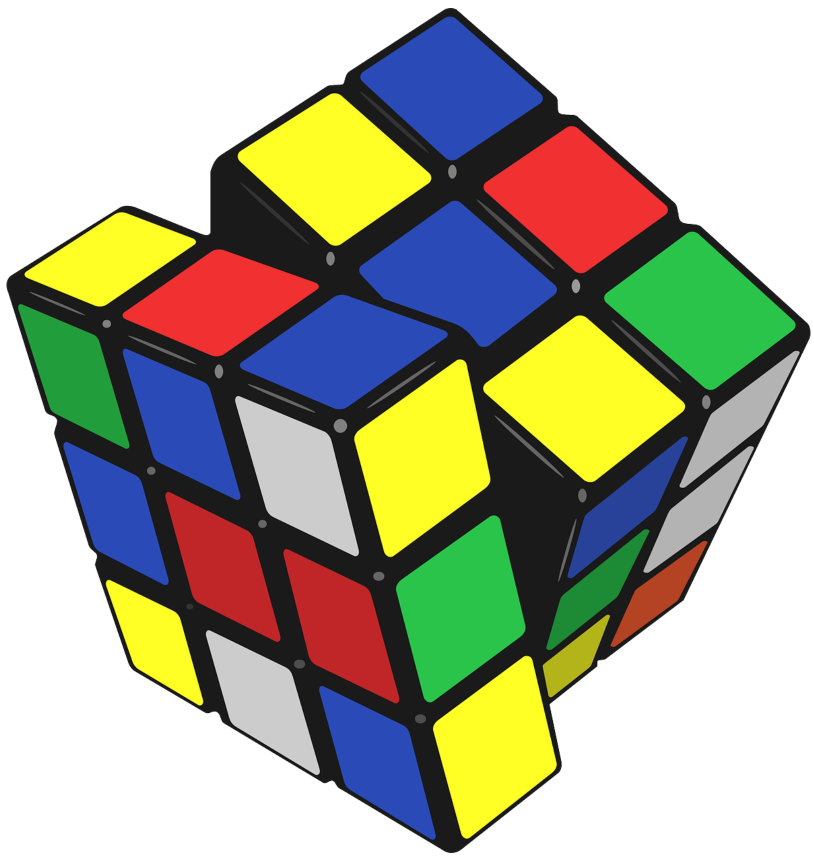


Figure 4.10: Sometimes accounting for all the possible arrangements of a situation is not easily accomplished with just the multiplication principle.

Considering this, we can count the arrangements by considering the four letters. Once we choose a letter to be first in a new arrangement, say, T, we have three letters left. If we choose the letter M to be the second in our arrangement, we now have two letters left. We can continue this way until we have used all four letters in a new arrangement. Each time we choose a letter, we have one less; this tells us the word MATH has $4 \cdot 3 \cdot 2 \cdot 1 = 24$ different configurations.

**Definition 4.2.6** This method of counting uses a **factorial**. We denote a factorial as $n!$ and it is calculated as

$$n! \;=\; n\;\cdot(n-1)\;\cdot(n-2)...3\cdot 2\cdot 1$$

Some special cases that won't fit this definition are:

- $0! = 1$
- $1! = 1$

■ **Example 4.24** Let's calculate 10!

$$10! = 10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1 = 3628800$$

Notice how we start multiplication with the number in the factorial and continue until we arrive at 1. ■

■ **Example 4.25** Let's see how the ratio between two factorials would work:

$$\frac{15!}{13!} = \frac{15\cdot 14\cdot 13\cdot 12\cdot 11\cdot 10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{13\cdot 12\cdot 11\cdot 10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1} = 15\cdot 14 = 210$$

■

Let's consider a slightly different type of scenario. What if we wanted to find how many different ways gold, silver, and bronze medals can be awarded if eight sprinters have made it to the Olympic finals in the 100-meter race?

In this scenario, seven sprinters remain once one of the sprinters wins the gold medal. When the next sprinter wins the silver medal, six sprinters remain one of whom could win the bronze medal.

We could use the Fundamental Counting Principle and find there are $8\cdot 7\cdot 6 = 336$ arrangements. However, in this scenario, the order of selection is essential, and we can generalize arranging, in order, $r$ items out of $n$ possibilities as **permutations**.

**Definition 4.2.7** A **permutation** is used to determine the number of possible arrangements in a set when the order of the arrangements matters. If we are selecting $r$ items out of $n$ possibilities, we use the formula.

$${}^{n}P_{r} = \frac{n!}{(n-r)!}$$

Notice that in a permutation, we consider how order factors into the situation. The notion of an order will allow us to identify that we must use a permutation.

■ **Example 4.26** Twenty-five people attend a charity benefit, and three gift certificates are given away as door prizes: one gift certificate is in the amount of \$100, the second is worth \$25, and the third is worth \$10. The prizes are worth different amounts, so order matters, and we use our permutation formula. Assuming no person receives more than one prize, how many different ways can the three gift certificates be awarded?

$$ {}^{25}P_3 = \frac{25!}{(25-3)!} = \frac{25!}{22!} = 25 \cdot 24 \cdot 23 = 13800 $$

■

■ **Example 4.27 — Pool Billiards.** You want to play billiards and care about the order of three balls for this particular game. In how many ways can you order 3 out of the 16 balls?

Since only the order of the three balls matter, we can use a permutation.

$$ {}^{13}P_3 = \frac{16!}{(16-3)!} = \frac{16!}{13!} = 16 \cdot 15 \cdot 14 = 3300 $$

■

■ **Example 4.28** How many ways can a four-person executive committee (president, vice-president, secretary, and treasurer) be selected from a 10-member board of directors of a non-profit organization? The order of selection is essential, so we calculate using a permutation.

$$ {}^{10}P_4 = \frac{10!}{(10-4)!} = \frac{10!}{6!} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040 $$

■

In the previous section, we considered the example of choosing four people out of a group of 10, where the order of selection was crucial. What if we wanted to know how many ways we could select a four-person executive committee from a 10-member board of directors of a non-profit organization? Notice in this revised situation that the order of choice is not important. When counting possible arrangements where order does not matter, we use **combinations**.

**Definition 4.2.8** A **combination** is used to determine the number of possible arrangements in a set when the order of arrangements does not matter. If we are selecting $r$ items out of $n$ possibilities, we use the formula.

$$^{n}C_{r} = \frac{n!}{(n-r)!r!}$$

■ **Example 4.29** How many ways is a four-person executive committee selected from a 10-member board of directors of a non-profit organization? This example differs from our previous example because the committee members do not have specific positions, so we use our combination formula.

$$^{10}C_{4} = \frac{10!}{(10-4)!4!} = \frac{10!}{6!4!} = \frac{10\cdot 9\cdot 8\cdot 7}{4\cdot 3\cdot 2\cdot 1} = 210$$

Notice we have fewer arrangements possible when order does not matter. ■

■ **Example 4.30** How many different ways can you choose three books from a summer reading book list of five books?

The book order does not matter, so we use our combination formula.

$$^{5}C_{3} = \frac{5!}{(5-3)!3!} = \frac{5!}{3!3!} = \frac{5\cdot 4\cdot 3}{3\cdot 2\cdot 1} = 10$$

■

■ **Example 4.31 — Pool Billiards revised.** We get back to our pool game, but this time we pick randomly three balls from the sixteen without caring about the order.

Since order doesn't matter, we can use the combination formula.

$$^{16}C_{3} = \frac{16!}{3!(16-3)!} = \frac{16!}{3!\cdot 13!} = 560$$

■

### 4.2.5 Counting Techniques and Probability

We can use our counting techniques to find probabilities of certain other types of events. One of the most famous problems involves finding the likelihood of sharing the same birthday with someone within a particular group of people.

■ **Example 4.32** Suppose three people are in a room. What is the probability that there is at least one shared birthday among them? There are many ways there could be at least one shared birthday, so we need to find an easier way to calculate this probability. Using the complement and finding the chance of no shared birthdays, we can simplify the problem using our counting techniques.

Because we know the probabilities of all possible outcomes must add up to 1, we can set up our calculation as

$$P(\text{at least one shared birthday}) = 1 - P(\text{no shared birthday})$$


Figure 4.11: We use the concept of the complement of a set to find the probability of the event and its opposite.

We start by computing the probability that there is no shared birthday. Imagine you are one of these three people. Your birthday can be any day of the year, so there are 365 choices for your birthday. What is the probability that the second person does not share your birthday? There are 365 days in the year (ignoring leap years), and by removing your birthday, there are 364 choices that will guarantee you do not share a birthday with this person, so the probability that the second person does not share your birthday is 364/365.

Now we move to the third person. What is the probability that this third person does not have the same birthday as you or the second person? Three hundred sixty-three days will not duplicate your birthday or the second person's, so the probability that the third person does not share a birthday with the first two is 363/365.

We want the second person not to share a birthday with you and the third person not to share a birthday with the first two, so we use the multiplication rule.

$$P(\text{no shared birthday}) - \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \approx\ 0.9918$$

Now subtract from 1, and our probability is

$$P(\text{at least one shared birthday}) = 1 - P(\text{no shared birthday}) = 1 - 0.9918 = 0.0082$$

The probability of at least one shared birthday among three people is relatively small. What if we increased the number of people and asked the same question? First, notice we could write the above problem as

$$P(\text{at least one shared birthday}) = 1 - \frac{{}^{365}P_3}{365^3}$$

■

■ **Example 4.33** Suppose 30 people are in a room. What is the probability that there is at least one shared birthday among these 30 people?

$$P(\text{at least one shared birthday}) = 1 - \frac{{}^{365}P_{30}}{365^{30}} \approx\ 0.706$$

Are you surprised that there is a 70% chance of at least one shared birthday among a group of 30 people? Problems like this in probability are not intuitive, but they are exciting and often surprising. ■

We can generalize this idea to independent events with the **at least once rule**.

**Definition 4.2.9 — At Least Once Rule for Independent Events**
Suppose the probability of an event $A$ occurring is $P(A)$. If we have $n$ attempts or opportunities for $A$ to happen then:

$$P(\text{At least one event A in } n \text{ trials}) = 1 - (P(\text{no A in 1 trial}))^n$$

■ **Example 4.34** The chance of getting at least one head in 30 coin tosses is

$$P(\text{At least one head in 30 tosses}) = 1 - (P(\text{not head}))^{30} = 1 - \left(\frac{1}{2}\right)^{30} = 0.9999$$

Note $P(\text{not head}) = P(\text{tails}) = \frac{1}{2}$ in one coin toss. ■

### 4.2.6 Exercises 4.2

For each of the following problems, express each of your probabilities as both fraction and a decimal rounded to three places.

**Problem 4.24** You draw two cards, with replacement, from a standard deck of cards. What is the probability both are red kings?

**Problem 4.25** You draw two cards, without replacement, from a standard deck of cards. What is the probability both are red kings?

**Problem 4.26** The probability that a person is afraid of thunderstorms is 35%. If we ask three random people, what is the probability all three are afraid of thunderstorms?

**Problem 4.27** In your drawer, you have ten pairs of socks, 6 of which are white. If you reach in and randomly grab two pairs of socks, what is the probability that both are white?

**Problem 4.28** A bag contains two green, four red, and six blue marbles. If you chose one marble randomly from the bag, what is the probability that it is a green or blue marble?

**Problem 4.29** If you select one card at random from a standard deck of cards, what is the probability it is a jack or a queen?

**Problem 4.30** If you select one card randomly from a standard deck of cards, what is the probability it is an ace or a heart?

**Problem 4.31** In a college math class of 20 students, there are 12 students majoring in criminal justice, 4 are first-year students, and 8 are sophomores. Eight political science majors, 2 are freshmen, and 6 are sophomores. If we select a student at random, what is the probability the student is either a freshman or a political science major?

**Problem 4.32** On a particular Saturday, the probability of rain is 54%, the likelihood the temperature will fall below 60°F degrees is 37%, and the chance both would occur is 17%.

(a) For this Saturday, find the probability that there will be rain or the temperature below 60°F.

(b) Suppose you decide not to go fishing this Saturday if it rained or if the temperature was below 50°F. What is the probability that you will go fishing?

**Problem 4.33** Jack has five different suits, six different shirts, ten different ties, and four different pairs of shoes. How many different outfits can Jack make if he wears a suit, a shirt, a tie, and a couple of shoes?

**Problem 4.34** Solve each of the following factorials:

(a) $5!$

(b) $(10-2)!$

(c) $\frac{5!}{2!}$

(d) $\frac{6!}{4!2!}$

**Problem 4.35** Calculate the following permutations.

(a) ${}^{15}P_2$

(b) ${}^{13}P_6$

(c) ${}^{4}P_1$

(d) ${}^{3}P_0$

**Problem 4.36** Calculate the following combinations.

(a) ${}^{23}C_6$

(b) ${}^{12}C_4$

(c) ${}^{3}C_0$

(d) ${}^{5}C_1$

**Problem 4.37** Alex's college course will allow him to choose from three foreign languages, three mathematics courses, five sciences, six physical education classes, four social sciences, and five English literature courses. How many different arrangements of his schedule are possible?

**Problem 4.38** There are 21 novels and 15 volumes of poetry on a reading list for a college English course. How many different ways can a student select one novel and one book of poetry to read during the semester?

**Problem 4.39** Eight sprinters have made it to the Olympic finals in the 100-meter race. How many ways can the gold, silver, and bronze medals be awarded?

**Problem 4.40** There are ten people in a civics club. In how many ways can four people be chosen as president, vice president, secretary, and treasurer?

**Problem 4.41** How many ways can we make a two-digit number from seven numbers if we can't repeat numbers?

**Problem 4.42** How many five-character passwords can be made using the letters A through Z if no repeated letters are allowed?

**Problem 4.43** How many ways can four members from a family of six line up for a photo session if order matters?

**Problem 4.44** How many ways can four members from a family of six line up for a photo session if the order doesn't matter?

**Problem 4.45** 12 candidates are running for any of 6 district positions in a school's student council. In how many different ways could the six positions be filled?

**Problem 4.46** Suppose five people are in a room. What is the probability of at least one shared birthday among these five people?

**Problem 4.47** Find the probability of drawing three queens in a row from a standard deck of cards when the drawn card is returned to the deck each time.

**Problem 4.48** Find the probability of drawing three queens in a row from a standard deck of cards when the drawn card is NOT returned to the deck each time.

**Problem 4.49** Find the probability of randomly drawing and immediately eating two red M&Ms in a row from a bag that contains eight red M&Ms out of 40 M&Ms total.

**Problem 4.50** Find the probability of:

(a) Getting a sum of either 10, 11, or 12 on a roll of two dice.

(b) Drawing either a red 6 or a black 8 on one draw from a regular deck.

**Problem 4.51** Find the probability of:

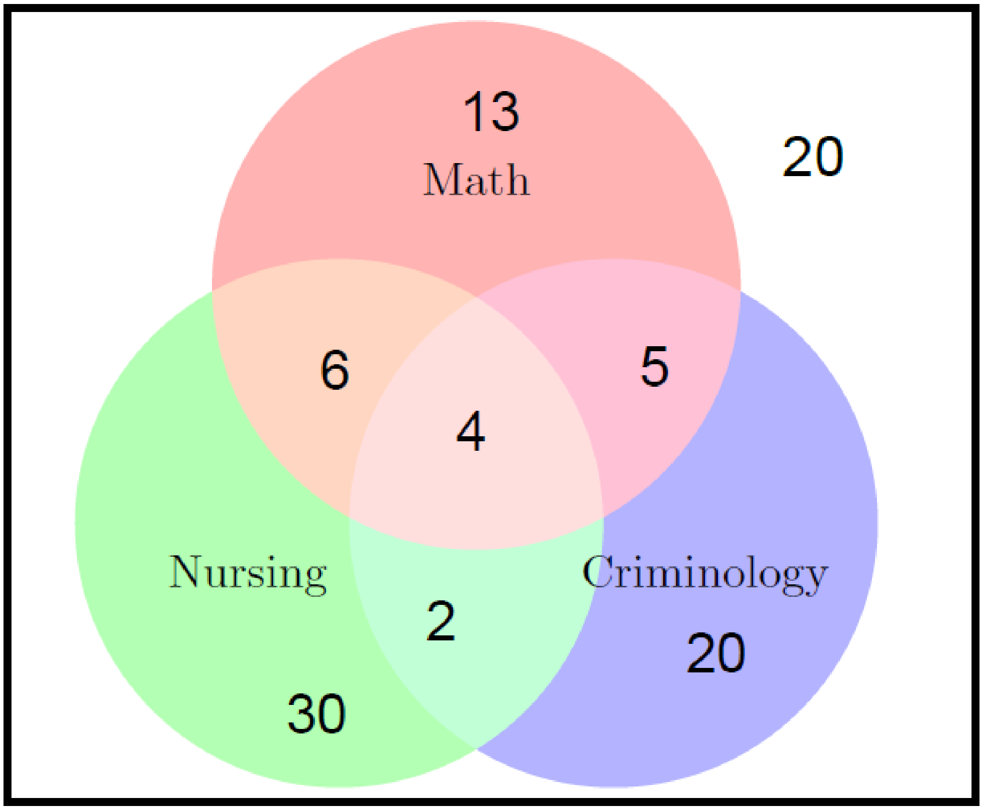
(a) Drawing either a jack or a spade from a regular deck of cards.

(b) Drawing either a face card (jack, queen, or king) or a diamond from a regular deck of cards.

**Problem 4.52** Find the probability of rolling at least one 6 in four rolls of a standard six-sided die.

**Problem 4.53** Find the probability of rolling at least one double-6 in 24 rolls of two six-sided dice.

**Problem 4.54** We interviewed 100 students about their class majors selection in the semester. From them, we collected the following information that is represented in the following Venn diagram.




If we use the Math circle as the event for students taking math classes. The Nursing circle is the event for taking nursing classes. And the Criminology circle is the event for taking Criminology classes. then calculate the following probabilities:

(a) The probability of a student taking classes from the three areas at once.

(b) The probability of a student not taking classes from any three areas.

(c) The probability of a student taking classes from Math and Nursing at once.

(d) The probability of a student taking classes from either Math or Nursing.

(e) The probability of a student taking classes from either Math or Nursing or Criminology.

(f) The probability of a student taking classes from either Math but none from Nursing or Criminology.

# 4.3 Project

## 4.3.1 Project:Gacha Games

Gacha games is the term used to refer to games where you gamble a specific currency for the chance of obtaining either characters or weapons for your campaign. Nowadays, in the video-game mobile market, many games have Gacha embedded mechanics to appeal to their audience with the gambling-type thrills they provide.

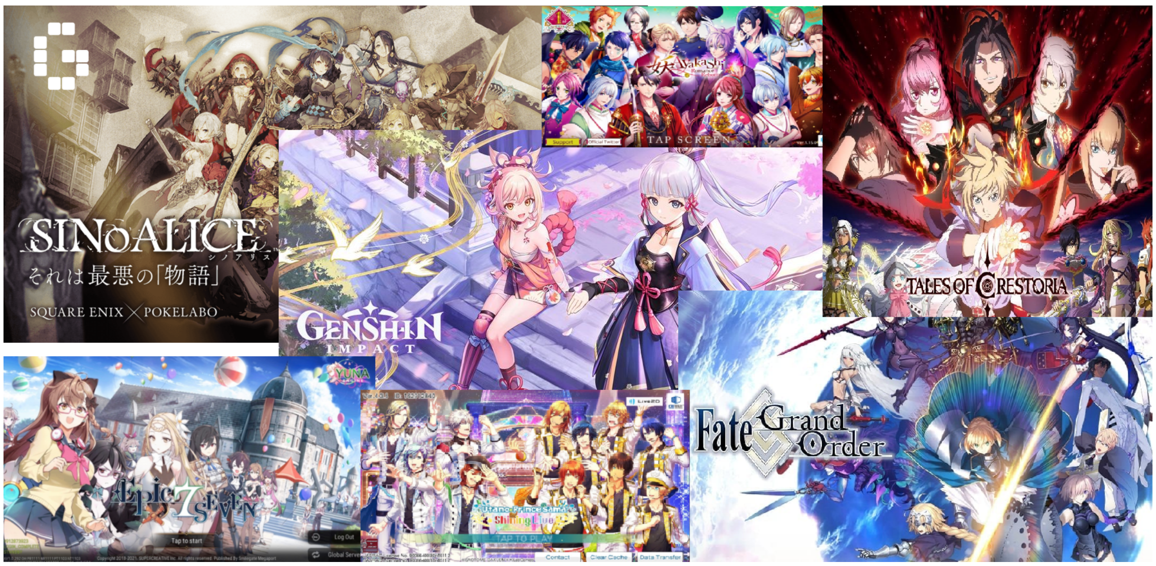



Figure 4.12: Featuring the following games: Sino Alice, Genshin Impact, Ayakashi,Tales of Crestoria, Epic Seven, Utano Princess Sama, Fate Grand Order.

**Objective:**

The goal of this project is to contrast experimental and theoretical probability.

**Instructions:**

(1). First, select a game that you want to use for this project.

(2). Collect information on the type of objects involved in the gacha mechanics, whether weapons or characters, and identify how the game's different tiers and rarity works.

(3). From the game information, obtain the theoretical probabilities from the objects by tier.

(4) Find a simulator for the particular game (Most games have simulators online for fun).

(5) Run the experiment of pulling for characters or weapons at least 100 times. (Often, gacha games do pull by ten objects).

(6) Register the frequency the different objects have and compare them with the theoretical probability.

(7) Run the experiment several times and check how much more is necessary so that the two probabilities get closer. (Law of large numbers).

**Example** Let's use Genshin Impact:

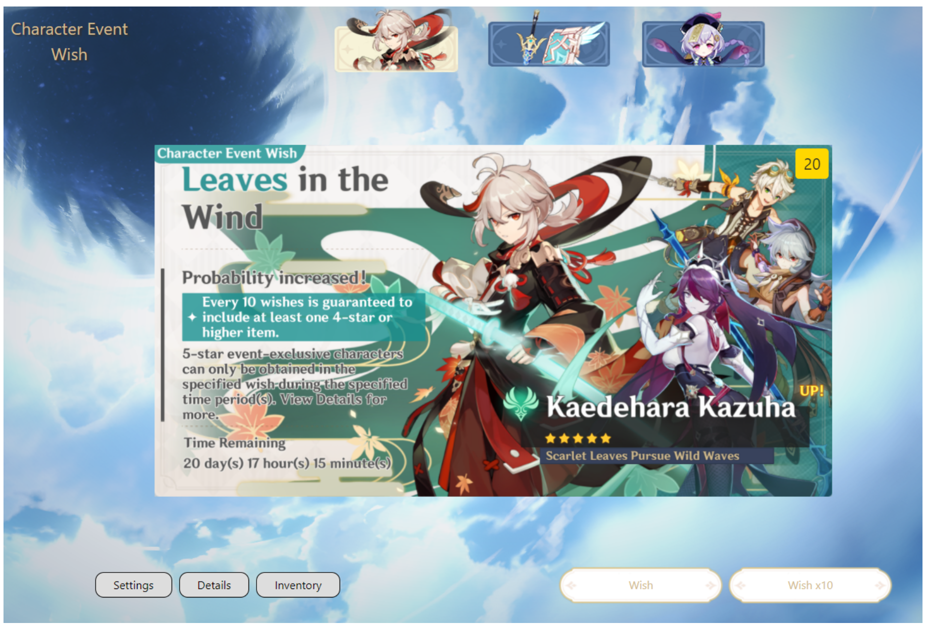


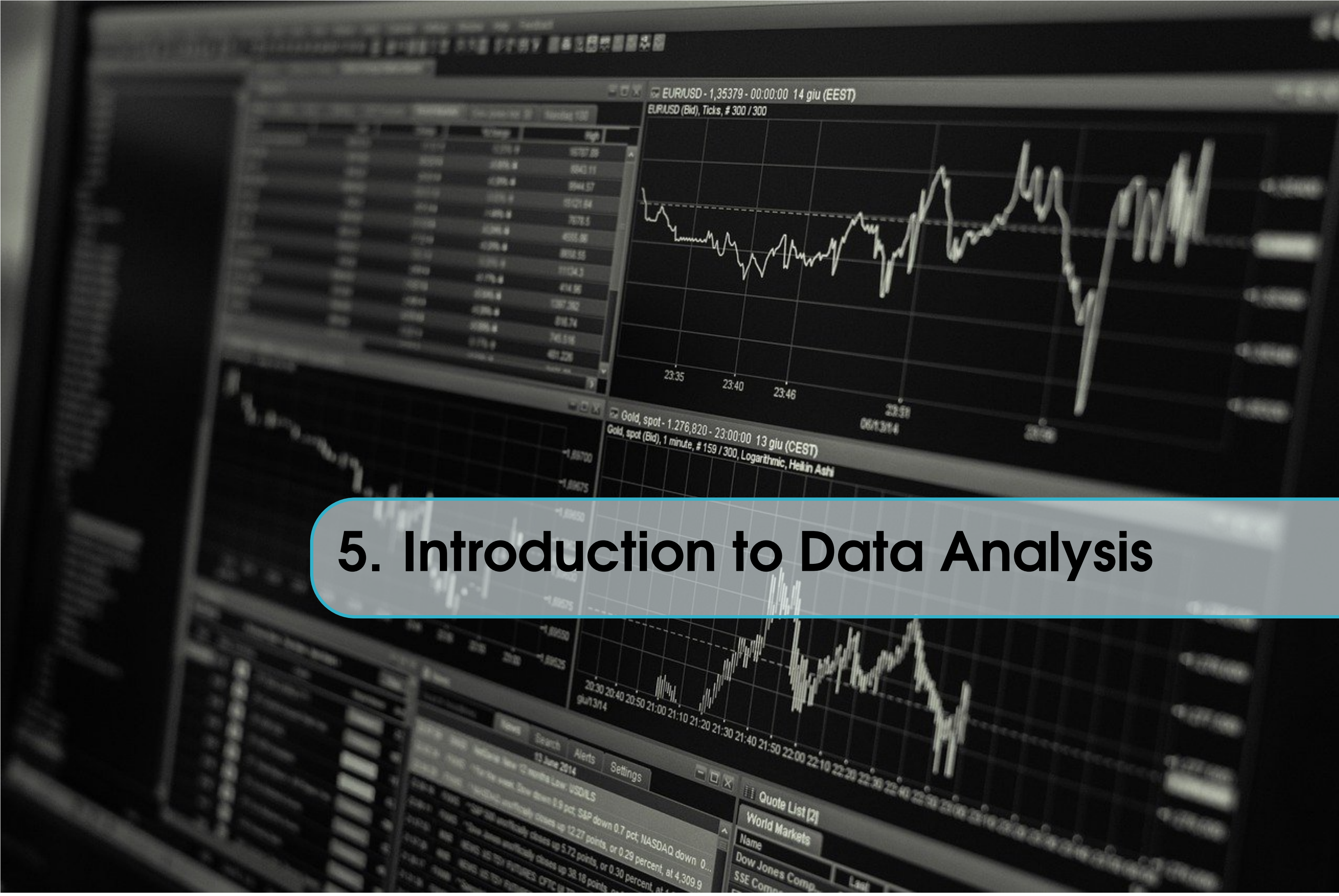

Figure 4.13: You can find the simulator for this game in `https://gi-wish-simulator.uzairashraf.dev/`

You can create a table to record the frequencies from every attempt (For example, the table below for three shots with ten wishes).

| Category | Pull 1 Frequency | Pull 2 Frequency | Pull 3 Frequency | Total |
|---|---|---|---|---|
| Weapon | | | | |
| 4 Starts Hero | | | | |
| Stars Hero | | | | |

and obtain theoretical probability from a website from the game.

`https://genshin-impact.fandom.com/wiki/Cumulative_Probability`




# 5. Introduction to Data Analysis

You may ask yourself, "When and where will I ever use statistics?" You will see statistical information if you read newspapers, watch television, or use the Internet. For example, there are statistics about crime, sports, education, the economy, politics, and real estate. It is essential to understand what the statistics mean and how you can use them to your advantage.

You will undoubtedly be given statistical information at some point in your life, and you need to know some techniques for analyzing the information thoughtfully. Think about personal events such as buying a house or managing a budget. Think about your chosen profession. Many fields of study, such as economics, business, psychology, education, biology, law, computer science, police science, and early childhood development, require at least one course in statistics. Statistics

This chapter addresses the basic ideas of statistics. You will also learn how data is collected and how to distinguish "good" from "bad" data. This chapter borrows from the Open Stax Introductory Statistics (see [Bar13]), Introductory Statistics from David Lane (see [Lan21]), and Introductory Statistics from Shafer and Zhang (see [SZ21]), but you can find a lot of these ideas in any book of elementary statistics.

## 5.1 Statistics Basics

The science of **statistics** deals with data collection, analysis, interpretation, and presentation. We produce, see and use data in our everyday lives.

In statistics, we generally want to study a population. You can think of a population as a collection of persons, things, or objects under study. To analyze the population, we select a sample. The sampling is to choose a portion (or subset) of the larger population and study that portion (the sample) to gain information about the population. Data are the result of sampling from a population.

**Definition 5.1.1 — Population and Sample**

The groups of people or objects we can study are limited, so it is important to create a clear definition.

- The **population** of a study is the entire group of things or people that are being studied.
- A **sample** of that population is a subset of the population that we choose for data gathering purposes.

Sampling is an efficient technique because examining an entire population takes time and money. Once we have data from a sample, we can make inferences about the behavior of the whole population.

■ **Example 5.1 — Grades at School** .

If you want to compute your school's overall grade point average, selecting a representative sample of students who attend the school will make sense. The data collected from the sample would be the students' grade point averages.

Figure 5.1: A classroom can be a sample for students attending your school.

■

■ **Example 5.2 — Presidential Elections** .

During political elections, opinion polls of samples of 1,000–2,000 people are collected to measure political climate and people's political inclinations.

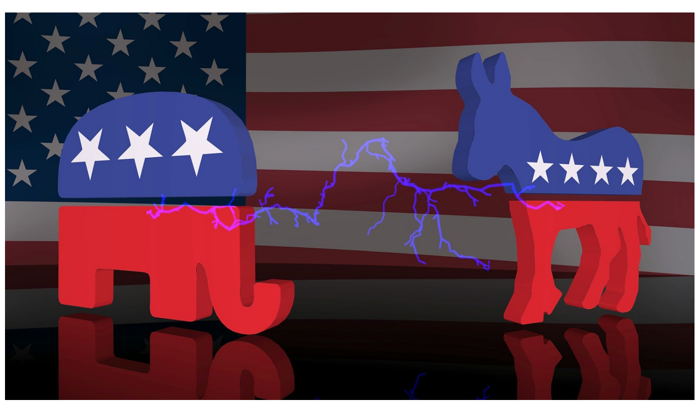


Figure 5.2: In good polls representations from all the parties involved is necessary.

Opinion polls should be designed to represent the views of the people in the entire country. ■

■ **Example 5.3 — Blood Samples**

Medical blood samples are an example of how we can measure general behavior using a small representative amount.

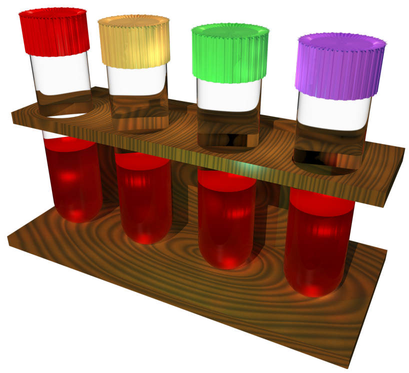


Figure 5.3: A small quantity of blood is used to measure the greater pool in our body.

Our blood taken from the correct places can accurately represent the blood in our body and, depending on what we measure, give important information about the measurements. ■

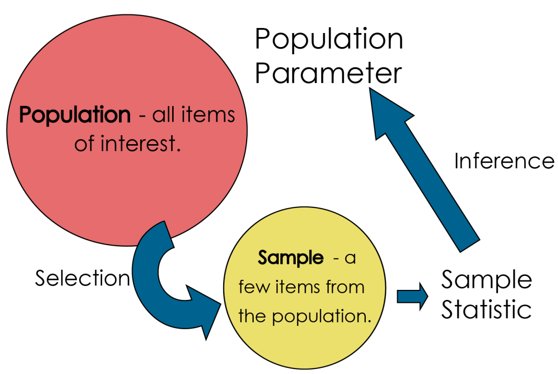
Once we understand that the sample presents us with a representation of a full-body (the population in this case), we can determine what we want to find out and how to measure it.

**Definition 5.1.2 — Parameter and Statistic**

Given a population from which we can obtain a sample, we define our measurements as follows:

- A **parameter** is a numerical value that summarizes the **Population** data.
- A **statistic** is a numerical value that summarizes the **sample** data.

We can summarize the process we want to achieve as follows:



The **sample statistic** is found to estimate or make an inference of the unknown population parameter. On the other hand, the **parameter** is a numerical characteristic of the whole population that a statistic can estimate.

**■ Example 5.4 — Math Class Grades**

Suppose you want to study the grades of the students at your school taking math classes.


Since hundreds of students are taking math classes, you will need a sample and then make inferences about the entire population. You could use one Quantitative Reasoning math class as a sample of the school students taking math classes. Let's identify each of these parts.

- **Population:** The entire population of students at the school taking math classes.
- **Sample:** The students enrolled in one Quantitative Reasoning math class.

We need to choose a variable that we can measure, such as the average number of points earned by students at the end of the term. Hence:

- **Population Parameter** The average number of points earned by students in math classes at the end of the term at the school.
- **Sample Statistic:** The average number of points earned by students in the Quantitative Reasoning math class at the end of the term.

■

■ **Example 5.5 — How to tell Statistics vs Parameters in statements?**

Given a statement with a measure, we can ask ourselves if we have a statistic or a parameter. With small populations, you usually have a **parameter** because the groups are small enough to measure:

- 40% of the US senators voted against the policy Z. We have 100 US Senators, and you can count what every single one voted; hence, this is a **parameter**.

- 60% of the graduates of a certain school got an SAT score above the national average. You can count every single one of the graduates of the school; hence, this is a **parameter**.

If the given fact is about a vast population, you likely have a **statistic**:

- The rate of unemployment in the US is the 15%. It is impossible to survey everyone in the US for their employment status, therefore, this is a **statistic**.

- 40% of the students in college in the US have a pet. Again, the group mentioned is a large population, we don't have means to survey every student, therefore, this is a **statistic**.

■

**■ Example 5.6 — Identify Population, Sample, Statistics and Parameter.**

A video-game company wants to measure the reception a new title will have in the European Market; for this, a closed beta is open with players from the UK, France, and Italy.


Lets identify each part for our analysis:

- **Population:** The European Market
- **Sample:** The closed beta with players from the UK, France and Italy.
- **Parameter:** The European Market reception
- **Statistic:** The reception of players in the closed beta

■

One of the main concerns in the field of statistics is how accurately a statistic estimates a parameter. The accuracy depends on how well the sample represents the population. The sample must contain the characteristics of the entire population to be a representative sample. We are interested in both the sample statistic and the population parameter in inferential statistics.

### 5.1.1 Choosing a Sample

Gathering information about an entire population often costs too much or is virtually impossible. Instead, we use a sample of the population. A sample should have the same characteristics as the population it is representing. Most statisticians use various methods of random sampling to achieve this goal. This section will describe a few of the most common forms.

**Definition 5.1.3 — Simple Random Sample.**

When using a **simple random sample**, any group of n individuals is equally likely to be chosen as any other group of n individuals. In other words, each sample of the same size has an equal chance of being selected.

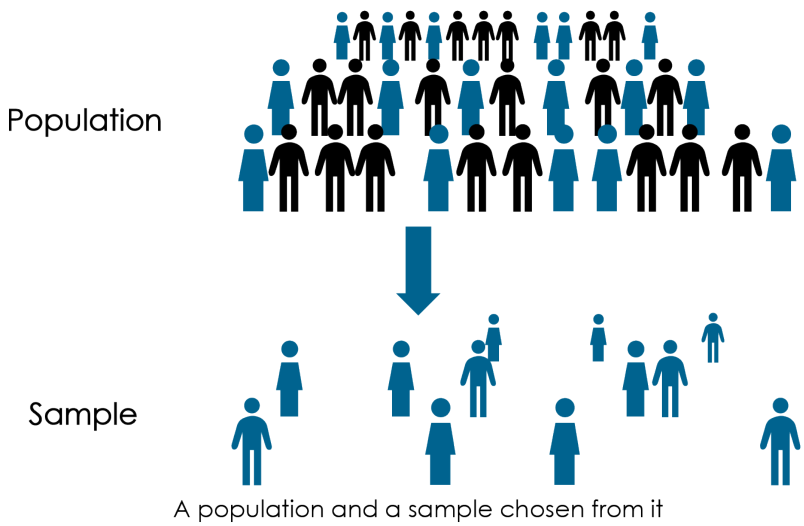



Figure 5.4: From a group of individuals we pick randomly members

**■ Example 5.7 — Simple Random Sample**

Consider the names of 25 students in a class are written on slips of paper and put in a hat. Then we select a sample of 10 names from the hat.

- Notice the selection process is randomized because we don't look inside the hat when making our selection.
- We can select the name of any of the 25 students in the class.

■

**Definition 5.1.4 — Systematic Sample**

**Systematic sampling** is a sampling technique where sample members are selected using a system to produce a fair selection. The system can be summarized as follows:

- We work with a population of $N$ elements, we select a random initial start.
- We define we want to obtain a sample of $n$ elements.
- We define the selection criteria to be $k = \frac{N}{n}$
- We select every $k^{th}$ element of population to obtain the sample.

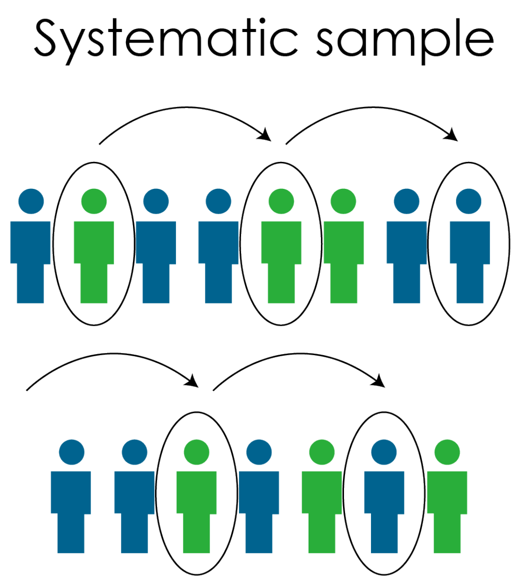



Figure 5.5: Selecting randomly an initial member and later using every third member in an arrangement is a system that can be used for systematic sampling.

**Example 5.8 — Examples of a Systematic Sample**

- A teacher gives a survey to every fifth student that comes to his classes.
- Every fourth customer from a Italian restaurant gets a satisfaction survey.
- From a population of 5000 farmers we extract a systematic sample of 10 farmers. (This means we select every $k^{th} = \frac{5000}{10} = 500$ name from a randomized list of the farmers).

■

**Definition 5.1.5 — Convenience Sample**

**Convenience sampling** is a technique where the sample is selected at the researcher's convenience. It is the most commonly used technique since it is an easy and inexpensive way to gather data.

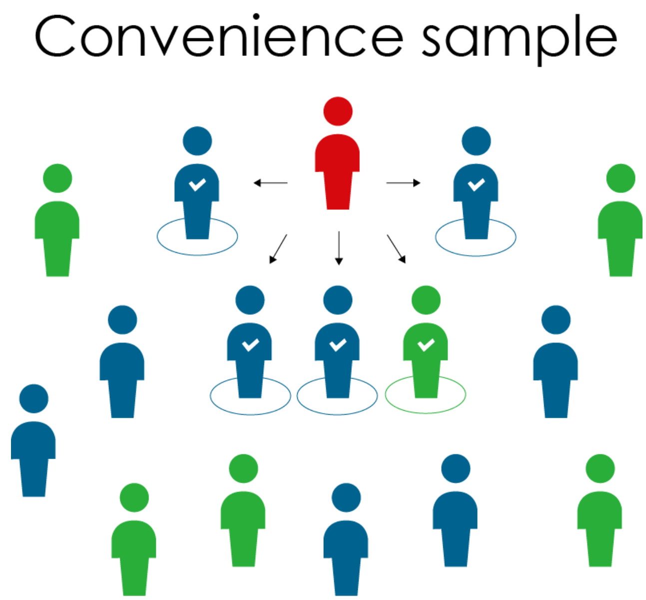



Figure 5.6: Convenience sampling is done by selecting the sample at the researcher's convenience in terms of accessibility. The researcher draws the sample from the part of the population that is close at hand.

**Example 5.9 — Examples of Convenience Samples**

A convenience sample is often selected from individuals that are conveniently available as:

- A speaker that selects the first 50 people to walk in the auditorium.
- Interviewing people who are downtown in your city.
- The front row seat students in a classroom.
- Cancer patients of your city.

Notice that convenience sampling is better used for pilot testing different elements and hypothesis generation but not for generalizing ideas. In this type of sampling, there is much left to the researcher's discretion, leading to bias and a lack of homogenization.

**Definition 5.1.6 — Stratified Sample**

This sampling technique involves the division of a population into smaller sub-groups known as strata, based on members' shared attributes or characteristics such as income or educational attainment.

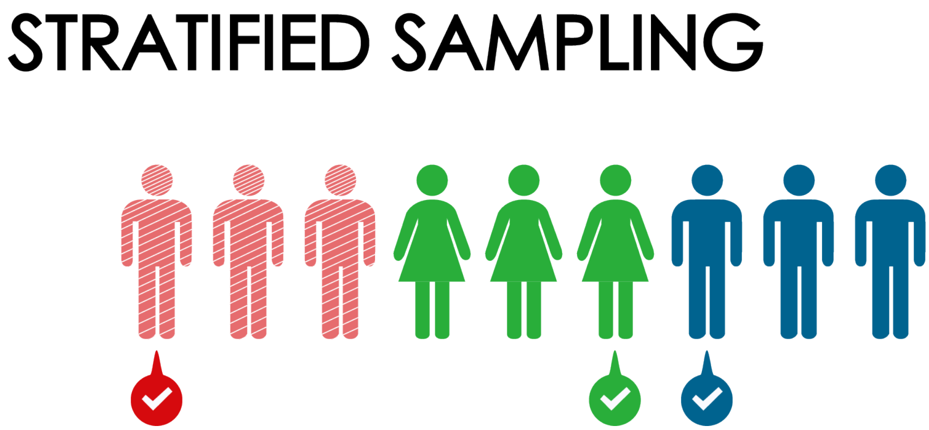



Figure 5.7: Stratified sampling can create different strata such as male and female, or separation by race, income, and other characteristics.

**■ Example 5.10 — Examples of a Stratified Sample**

- A company has 150 female employees and 125 male employees. They want to ensure that the sample reflects the gender balance of the company, so they sort the population into two strata (groups) based on gender. Then they use random sampling in each group, selecting 55 **women** and 45 **men**, which gives them a representative sample of 100 people.

- The government wants to survey political inclinations among people of diverse ages. They define age brackets as: **young adult** from 18 to 30 years old, **adult** from 31 to 50 years old, and **senior** from 51 and up. The firm in charge obtains a representative sample of 100 people in each of the age brackets.

- A public health department runs a study measuring cholesterol levels at certain ages and designs a stratified sample by race, with 50 volunteers in each of the strata for **Asian, Latin, White, Afro-Americans** and **Pacific Islanders**.

- A Real State Company designs a stratified sample of their customers, aiming to study different preferences in the housing environment. The strata is create based on household income: **Low-Range** is 100,000 dollars per year and less, **Mid Range** is 100,000-400,000 dollars per year, and **High Range** is above 400,000 per year.

■

**Definition 5.1.7 — Cluster Sampling**

The cluster sampling technique splits a population into smaller groups known as clusters. Then randomly selected whole clusters form the sample.

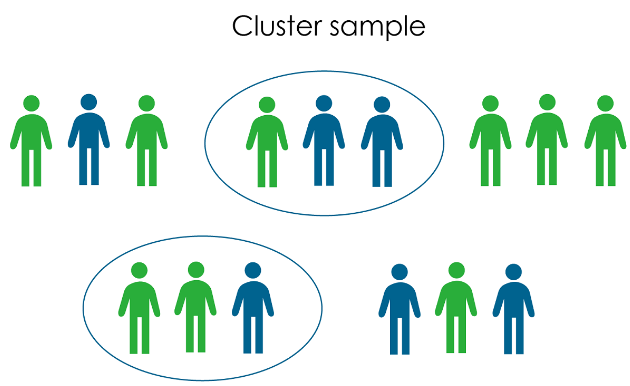



Figure 5.8: We divide our population into clusters of three people. Then we can randomly select clusters of three to create our sample.

**■ Example 5.11 — Example of a Cluster Sample**

A researcher may be interested in data about city taxes in Georgia. The researcher would get data from selected cities and compile them to get a picture of the state. The individual cities would be the clusters in this case. ■

Clusters can be confused with strata. To better understand the difference between a cluster sample and a stratified sample, look at figure 5.1.1.

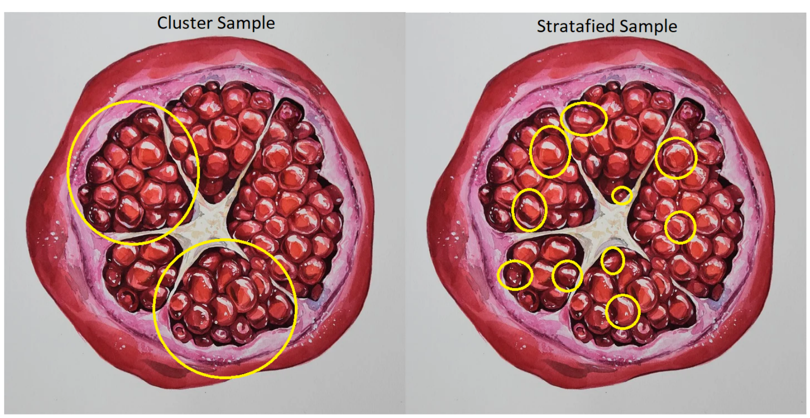
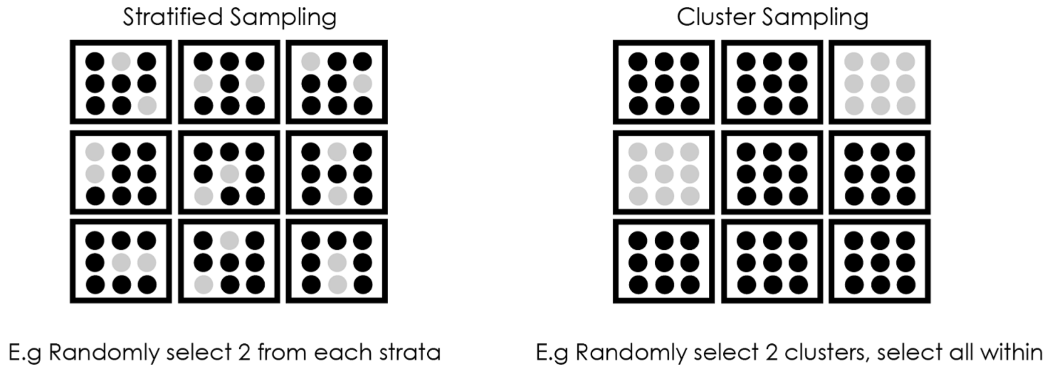



Figure 5.9: The picture on the left shows a cluster sample. Two entire clusters have been randomly chosen to be the sample. The picture on the right shows a stratified sample. Here, two seeds were randomly chosen from each of the clusters to be the sample.

■ **Example 5.12 — Another Cluster vs Strata Sample**

Consider a population of 81 people, we can obtain a sample in two ways:




The gray dots in each case are the samples selected. ■

■ **Example 5.13 — GGC sampling**

A study is done to determine the average tuition that Georgia Gwinnett College students pay per semester. Each student in the following samples is asked how much tuition they paid for the Fall semester. What is the type of sampling in each case?

(a) A sample of 100 undergraduate Georgia Gwinnett College students is taken by organizing the students' names by classification (freshman, sophomore, junior, or senior) and then selecting 25 students from each.

(b) From an alphabetical list of all undergraduate students in the Fall, every fiftieth student is chosen until 150 students are included in the sample.

(c) A random number generator is used to get a sample of 150 student identification numbers. Each group of 150 undergraduate students in the fall semester has the same probability of being chosen at any stage of the sampling process.

(d) The freshman, sophomore, junior, and senior years are numbered one, two, three, and four, respectively. A random number generator is used to pick two of those years. All students in those two years are in the sample.

(e) An administrative assistant is asked to stand in front of the library one Wednesday and ask the first 100 undergraduate students he encounters what they paid for tuition for the Fall semester. Those 100 students are the sample.

**Solution:** (a). Stratified (b). Systematic (c). Simple Random (d). Cluster (e). Convenience

■

### 5.1.2 Types of Statistical Studies

There are two main types of statistical studies: experiments and observational studies.

Figure 5.10: Experiments include more than those done in labs.

**Definition 5.1.8 — Experimental Study**

An **experimental study** is a statistical study in which a treatment is given to the sample members. In other words, something is done to change the environment of the sample members.

In an experimental study, we often have a control group and a treatment group:

- The **control group** is the group of sample members who do not receive the treatment.
- The **treatment group** is the group of sample members who do receive the treatment.

**Example 5.14 — The Clinic Experiment**

A common example of this type of study is a medication experiment. A pharmaceutical company wants to know if a new allergy medicine is effective. Of the study population, 500 were chosen as the sample members.

- Half of the sample, the treatment group, will be given the medicine.
- Half of the sample, the control group, will receive a fake pill, or placebo, that looks like the medicine.

■

**Definition 5.1.9 — The Placebo Effect**

The **placebo effect** is an effect produced by taking a placebo drug or treatment, which cannot be attributed to the properties of the placebo itself. It must therefore be due to a patient's belief in that treatment.

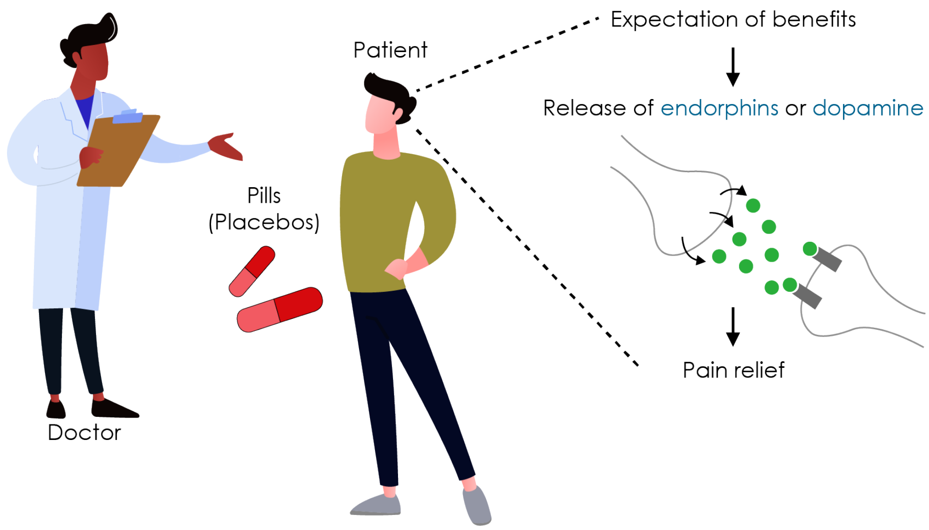



Figure 5.11: How the placebo effect works.

It is important to note that a "placebo" and the "placebo effect" are different things. The term placebo refers to the inactive substance itself. In contrast, the term placebo effect refers to any impact of taking a medicine that cannot be attributed to the treatment itself.

By having a control group and a treatment group, some people in each group will have the placebo effect. Therefore, if the results are better for the treatment group, researchers can conclude the medication is effective.

**Definition 5.1.10 — Blinding**

An experiment is said to be **single-blind** if the participants do not know whether they are in the control group or the treatment group.

An experiment is said to be **double-blind** if neither the participants nor the researchers know who is in the control group and who is in the treatment group.

In an experiment, researchers get better results if the study is blinded. If a person knows whether or not they are getting the treatment, it can change how they react to it. Also, if the researchers know, they may inadvertently ask questions differently when interviewing a participant. When interviewing a participant, the researcher will not have any reason to lead the participant toward their expected answers. Instead, each participant would have an identification number based on which medicine they received, identifying the group they belong to when the experiment is over.

**Definition 5.1.11 — Observational Study**

An **observational study** is a statistical study in which the researcher is merely observing and gathering data but does not have any influence or effect on the sample members.

Figure 5.12: In an observational study, we are only witnesses and recorders of behaviors or occurrences we want to study.

**Example 5.15 — Heart Disease in Smokers**

What if researchers wanted to know if smokers have a higher risk of heart disease? They might pick a random sample of 100 smokers and survey them about their health.

In this example, the researchers did not change the participants' behavior. They did not give them a new kind of cigarette, medication, or other change. They merely analyzed the data collected from the 100 participants. ∎

Why did the researchers decide to do that study as an observational study instead of an experiment? They would have asked people to smoke if they had experimented by comparing smokers to non-smokers; this is not morally justifiable since we have evidence that smoking harms your health. Surveying people who already smoke does not have to worry about the moral and ethical implications.

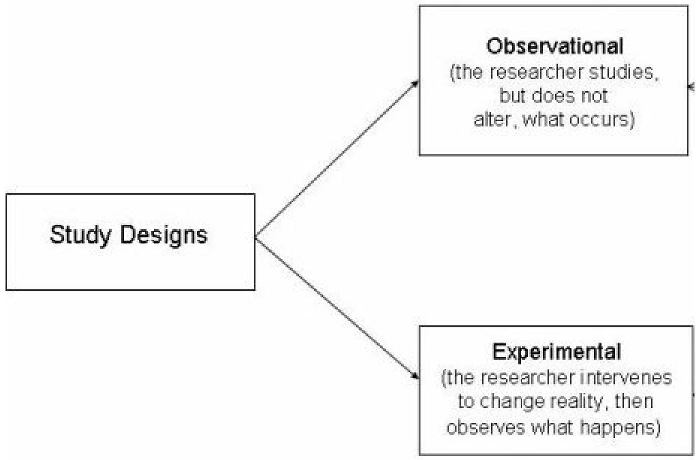



Figure 5.13: Types of Statistical Studies

Given a situation where we want to design a study, we would like to determine whether the study should be an experimental or an observational study.

**■ Example 5.16 — Observational or Experimental?**

Classify each of the following cases:

(a) Consider someone on the busy street of a New York neighborhood asking random people that pass by how many pets they have, then taking this data and using it to decide if there should be more pet food stores in that area.

(b) A psychologist wants to know if music helps students concentrate. The students take their exams. Half of them are in a room with music, and half are in a room without music.

(c) A insurance company wants to determine the efficiency of security measures in cars of a certain brand. For that, they compare the record of accidents from 1980 to 2020 of cars of the brand, observing accidents where the passengers used the security measures (example: seat belts) versus where they did not.

**Solution:** (a). An observational study (b). An experimental study (c). Observational Study (This is called a **case control study**. We can be confused by the existence of two groups, but note researchers apply no treatment, hence no experiment.) ■

### 5.1.3 Types of Data

Once you have decided the type of study you want to do, how to select your sample, and what you want to measure, the next step is to manage the data you will obtain from the study. We need to understand the different types of data our studies can produce. Mainly, there are two types of data that we use in statistics.

**Definition 5.1.12 — Qualitative and Quantitative data**

The data obtained from a study falls into one of two categories:

- **Qualitative data** is data based on categorizing or describing attributes of a population. Qualitative data are also known as **categorical data**. This kind of data is obtained via observed phenomenon, not through measurement, and is generally given labels but can be assigned numerical values.

- **Quantitative data** are the result of counting or measuring attributes of a population. In general, it is information about quantities that we can measure, so we describe them in terms of numbers. As such, it is also called **numerical data**.

**Example 5.17 — Examples of Qualitative Data**

**Qualitative data** includes observations we can make about our sample which help us identify different properties and attributes. For example:

- Hair color
- Ethnic group
- Gender
- Blood type
- Types of cars
- Eye color

■

### ■ Example 5.18 — Examples of Quantitative Data

**Quantitative data** is measured data using systems and tools for measuring. The units of measure are an important factor of this kind of data. For example:


Figure 5.14: Some tools for measuring blood sugar levels.

- Amount of money
- Pulse rate
- Weight
- Time
- Number of students in a class
- Height
- Birth rates in a town
- Average GPA of a student

■

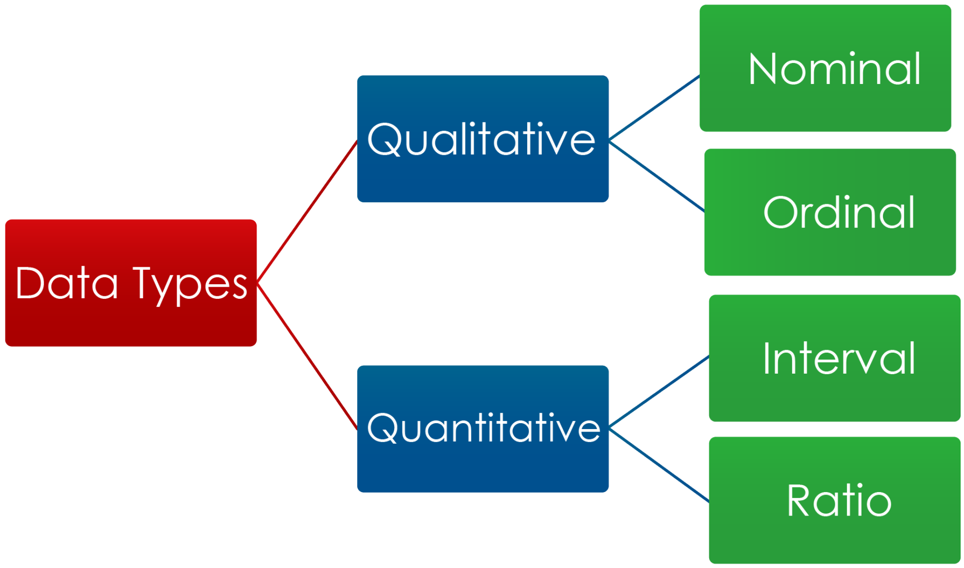
We can further express These two types of data in different forms.

For **qualitative data**, we can classify data as:

- Nominal data
- Ordinal data

For **quantitative data**, we can classify data as:

- Interval
- Ratio




**Nominal data** are essentially named variables, whereas **ordinal data** is called and ordered in some way.

■ **Example 5.19** The following variables are examples of nominal data:

- Gender
- Political preference
- Video-game console you own

The following variables are examples of ordinal data:

- Satisfaction levels
- Position on the Dean's list
- Grade in a class

Note that the nominal variables will store values as male, female, Republican, Democrat, Xbox, Nintendo, and PlayStation. The ordinal variables will store values as: Very Satisfied, Neutral, Unsatisfied, or First, Second, Third, or A, B, C. All of the ordinal variables have an order or tier. ■

**Ratio data** are essentially numbers that provide an accurate measure of the variable. They are calculated by assuming that the variables have an option for zero. On the other hand, **interval data** is where we can only produce the measure within a range of values.

■ **Example 5.20** The following variables are examples of ratio data:

- Height
- Weight

The following variables are examples of interval data:

- Family income
- Temperature averages in your city

The ratio variables will store values like 150 pounds or 175 centimeters, while the interval variables will store $[85000, 110000]$ dollars per year or $[5, 40]$ degrees Celsius. ■

■ **Example 5.21** Decide what types of data describes each data set below.

1. The number of pairs of shoes you own.
2. The type of car you drive.
3. The distance it is from your home to the nearest grocery store.
4. The number of classes you take per school year.
5. The type of calculator you use.
6. Weights of sumo wrestlers.
7. Number of correct answers on a quiz.
8. Post Traumatic Stress Disorder level.
9. Finishing place in a marathon race.

**Solution:**

| | | |
|---|---|---|
| 1. Quantitative, Ratio. | 4. Quantitative, Interval. | 7. Quantitative, Ratio. |
| 2. Qualitative, Nominal. | 5. Qualitative, Nominal. | 8. Qualitative, Ordinal. |
| 3. Quantitative, Ratio. | 6. Quantitative, Interval. | 9. Qualitative, Ordinal. |

■

### 5.1.4 Exercises 5.1

**Problem 5.1** In each case, identify Population, Sample, Statistic, and Parameter.

(a). A fitness center is interested in the mean of the time a client exercises in the center each week.

(b). Ski resorts are interested in the mean age that children take their first ski and snowboard lessons. They need this information to plan their ski classes optimally.

(c). A cardiologist is interested in the mean recovery period of her patients who have had heart attacks.

(d). Insurance companies are interested in the mean health costs each year of their clients so that they can determine the costs of health insurance.

(e). A politician is interested in the proportion of voters in his district who think he is doing a good job.

(f). A marriage counselor is interested in the proportion of clients she counsels who stay married.

(g). Political pollsters may be interested in the proportion of people who will vote for a particular cause.

(h). A marketing company is interested in the proportion of people who will buy a particular product.

**Problem 5.2** Use the following information to answer the following questions: A Georgia Gwinnett College instructor is interested in the mean number of days GGC math students are absent from class during a semester.

- What is the population she is interested in?

  (a). All GGC students.

  (b). All GGC English students.

  (c). All GGC students in her classes.

  (d). All GGC math students.

- The instructor's sample produces a mean number of days absent of 3.5 days. This value is an example of a:

  (a). Parameter

  (b). Data

  (c). Statistic

  (d). Population

- The instructor would like to find the population parameter. Find the true statement:

  (a). The instructor cannot know the exact value unless her data is from every student in the population.

  (b). The instructor knows it will be equal to the sample statistic.

  (c). The instructor knows it will be greater than the sample statistic.

  (d). The instructor knows it will be less than the sample statistic.

**Problem 5.3** Classify the following data variables as qualitative or quantitative. Include if they are nominal, ordinal, interval or ratio.

(a). The number of tickets sold to a concert.

(b). Percent of body fat.

(c). Favorite baseball team.

(d). Time in line to buy groceries.

(e). Most watched You Tube video.

(f). Most popular brand of toothpaste.

(g). Distance to the closest movie theater.

(h). Age of executives in Fortune 500 companies.

**Problem 5.4** State what kind of sampling method is used for each of the following:

(a). The instructor takes her sample by gathering data on five randomly selected students from each GGC math class.

(b). A study on users of a local park in Lawrenceville was interested in the visitor's age and frequency of use. A sample was randomly selected in a neighborhood with the first house and then every eighth area house.

(c). A woman in the airport is handing out questionnaires to travelers asking them to evaluate its service. She does not ask travelers who hurry through the airport with their hands full of luggage but instead invites all travelers near gates not to nap while they wait.

(d). The librarian at a public library wants to determine what proportion of the library users are children. She records this data for every fourth patron who checks out books.

(e). The marketing manager for an electronics chain store wants information about the ages of its customers. At each store location, around 100 customers fill out a survey that asks for this data and other variables of interest.

(f). A study is realized on the time students need to study on campus. A computer's random number generator randomly chooses five hundred student identification numbers. These students are then surveyed.

**Problem 5.5** Decide if each is an experiment or an observational study.

(a). A zoologist wants to understand the behavior of mountain lions. She sends out drones with cameras to video a region with mountain lions and then analyzes the data.

(b). A doctor wants to know if smokers or non-smokers are more likely to survive past 60.

(c). A psychiatrist wants to test a new anti-depressant. He gave 150 participants the anti-depressant to take every day and gave another 150 participants a placebo to take every day.

(d). A professor wants to know if calming music helps students during tests. She breaks her classes in half and has half of the students take a test in a quiet room while the other students take the same test in a room with calming music.

(e). A study at the University of Southern California separated 108 volunteers into groups based on psychological tests designed to determine how often they lied and cheated. Those who tended to lie had a different brain structure than those who did not lie (British Journal of Psychiatry).

(f). A double-blind drug versus placebo study of 103 patients suffering from tinnitus (the perception of ringing in the ears) demonstrated the effectiveness of Ginkgo Biloba extract. The ginkgo treatment improved the tinnitus patients' condition (Annals of Otology, Rhinology, and Laryngology).

**Problem 5.6** For the following research studies, describe the population, sample, population parameters, and sample statistics:

(a). Anthropologists determine the average size of early Neanderthals in Europe by studying skulls found at three sites in southern Europe.

(b). Astronomers typically determine the distance to a galaxy (a huge collection of billions of stars) by measuring the distances to just a few stars within it and taking the mean (average) measurements.

(c). The Higher Education Research Institute conducts an annual study of attitudes of incoming college students by surveying approximately 241,000 first-year students at 340 colleges and universities. There are approximately 1.4 million first-year college students in this country.

**Problem 5.7** You want to determine the typical dietary habits of students at a college. Which of the following would make the best sample and why? Also, explain why each of the other choices would not make a good sample for this study.

(a). Students in a single dormitory.

(b). Students in public health.

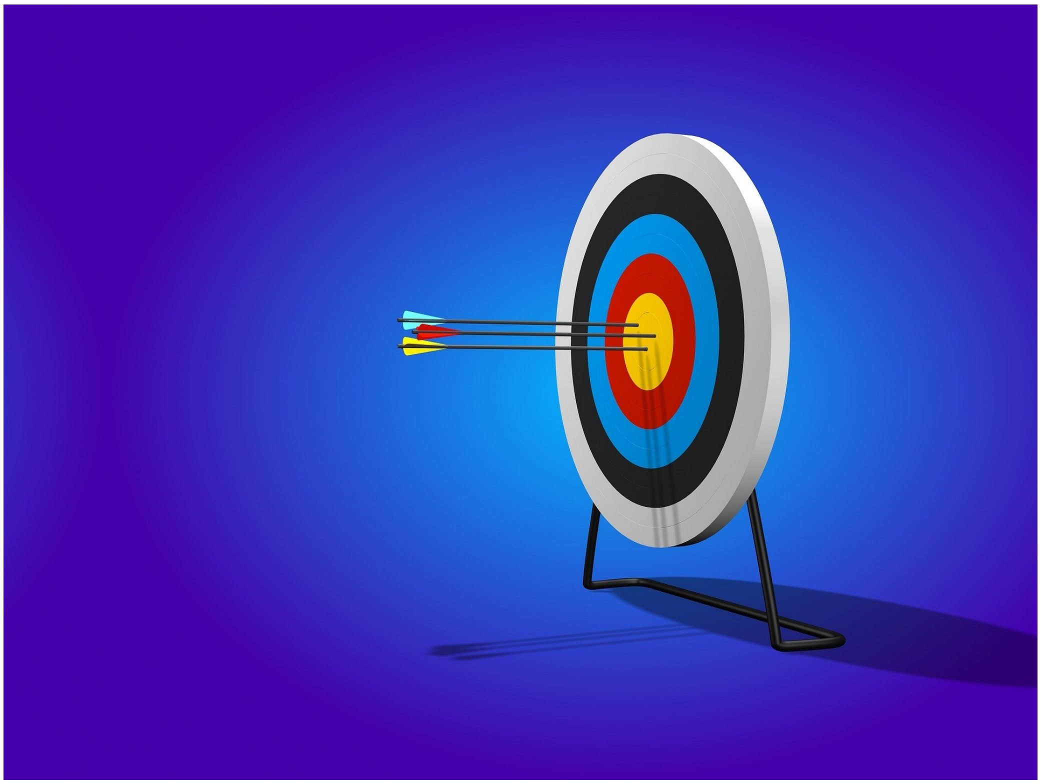
(c). Students in collegiate sports team.

(d). Students enrolled in a math class.

## 5.2 Descriptive Statistics

**Descriptive statistics** allows you to take large amounts of data and organize and summarize it. The statistics summarize particular data characteristics, often with a single number.

### 5.2.1 Measures of Central Tendency



A **measure of central tendency** is a single value describing a data set by identifying the central position. Measures of central tendency are sometimes called measures of central location. They are also called summary statistics. The mean (often called the numerical average) is the most common measure of central tendency, but there are others, such as the median and the mode.

**Definition 5.2.1 — Mean** .

The mean is equal to the sum of all the values in the data set divided by the number of values in the data set.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

**■ Example 5.22 — Sample set**

Find the mean of the sample set:

$$\{2,3,6,7,8\}$$

We identify each of the numbers in the sample as a data point $a_i$, or

$$a_1 = 2, a_2 = 3, a_3 = 6, a_4 = 7, a_5 = 8.$$

using the formula above we have the mean is:

$$\bar{x} = \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{(2+3+6+7+8)}{5} = 5.2$$

■

**Definition 5.2.2 — The mean with frequency**

Sometimes a number appears multiple times within a set of data. When this happens, instead of writing the same number multiple times in the mean calculation. We can write the number once and then multiply by the number of times it appears (also called **frequency**), and then divide the total sum of the frequencies (which will amount to the total of data points).

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i \times f}{\sum f}$$

**■ Example 5.23 — Mean with frequency in a larger data set**

The following data is based on responses to a survey question of how many pets people have in their household:

$$\{0,0,0,1,1,1,1,1,1,2,2,2,2,2,3,3,4,4,4,4\}$$

We can create a summary of the data and the frequencies:

| | | | | | |
|---|---|---|---|---|---|
| f | 3 | 6 | 5 | 2 | 4 |
| x | 0 | 1 | 2 | 3 | 4 |

Frequency tables are very useful for organizing data. Let's calculate the mean based on information:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i \times f}{\sum f}$$

$$= \frac{(0\times 3 + 1\times 6 + 2\times 5 + 3\times 2 + 4\times 4)}{(3+6+5+2+4)} = \frac{38}{20} = 1.9$$

■

Notice in both examples that the mean is not a whole number. Does that indicate there is a house with 1.9 pets? Of course not! The mean only tells you the numerical average of the data.

**Definition 5.2.3** The **median** is the middle number for a set of data that has been arranged from smallest to largest.

Note: If there are an odd number of data points, there is one middle number. If there are an even number of data points, you find the mean of the two middle numbers.

**Example 5.24 — Median of a data set with odd number of values**

Find the median of the elements of the set:

$$\{13, 24, 15, 37, 29\}$$

**Solution:**

To find the middle number, we must first put the numbers in ascending order:

$$\{13, 15, 24, 29, 37\}$$

The median (middle) value is 24.

**Example 5.25 — Median of a data set with an even number of values**

Find the median of the sample set:

$$\{98, 23, 7, 4, 22, 89, 16, 52\}$$

**Solution:**

To find the middle number, we must first put the numbers in ascending order:

$$\{4, 7, 16, 22, 23, 52, 89, 98\}$$

Because we have an even number of values (you can count eight), there is no one number in the middle. However, we take the two most central numbers, 22 and 23, and find their mean:

$$\frac{22+23}{2} = 22.5$$

So the median = 22.5

**Definition 5.2.4** The **mode** is the most frequent value(s) in our data set.

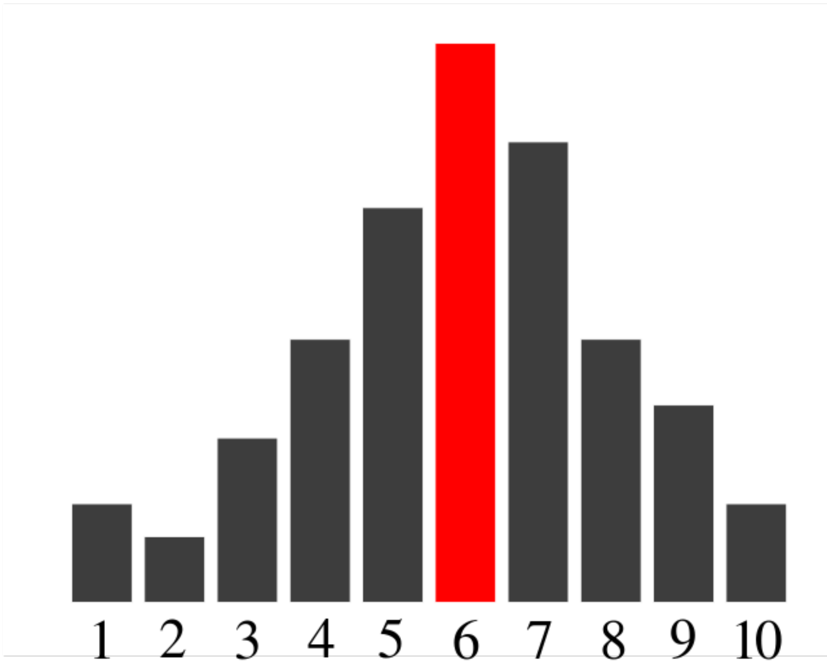



Figure 5.15: In visual representations, the mode will always stand out.

■ **Example 5.26 — Math test scores**

The Math test scores for 10 students are recorded as follows

$$\{72, 72, 72, 76, 78, 81, 83, 84, 90, 93\}$$

Find the median and mode.

**Solution:**

Notice the sample data is already ordered. This matters for our median, which is calculated as the average of the two middle numbers:

$$\frac{78+81}{2} = 79.5$$

For the mode, look for the most frequent score, 72, and it occurs thrice. Notice that the mode can be found whether the sample is ordered or not.

■

Data is often prone to repeated values and hence is considered a mode. However, we can have data sets with no mode at all. Consider the following example:

**Example 5.27 — Multi-modal and no-mode data sets**

Find the mode for each data set below:

(a) $\{1,4,5,5,8,10,10,12\}$

(b) $\{1,5,8,9,12,15,17\}$

**Solution:**

(a) Both 5 and 10 occur most often, so both 5 and 10 are modes.

(b) In this set, there is no number that repeats, so there is no mode.

**Definition 5.2.5** An **outlier** is a value in the data set which is abnormally bigger or abnormally smaller compared to the other data values.

Figure 5.16: An outlier always stands out from the other values.

■ **Example 5.28** Consider the data set

$$\{1,2,5,8,32\}$$

Are there any outliers?

**Solution:**

Notice the number 32 is much larger compared to other numbers. Only 2 or 3 units separate the other numbers. Considering that 32 is abnormally more prominent than the other values, we can consider 32 an outlier. ■

### How Outliers Affect the Mean, Median, and Mode

Because outliers are nowhere near the middle of the data set and do not repeat, outliers do not have any effect on the median or the mode. However, having an outlier pulls the mean in the direction of the outlier. In other words, a significant outlier makes the mean larger. A small outlier makes the mean smaller.

■ **Example 5.29 — Skewing the mean** .

The following three data sets are essentially the same except for the last two, where there is an outlier.

- $A = \{37,41,50,58,62,65\}$
- $B = \{1,37,41,50,58,62,65\}$
- $C = \{37,41,50,58,62,65,245\}$

Find the mean of all three data sets.

**Solution:**

For set A, the mean is

$$\bar{x} = \frac{37+41+50+58+62+65}{6} = 52.2$$

For set B, the mean is

$$\bar{x} = \frac{1+37+41+50+58+62+65}{7} = 44.9$$

For set C, the mean is

$$\bar{x} = \frac{37+41+50+58+62+65+245}{7} = 79.7$$

Notice in set B, the small outlier, 1, lowered the mean below that of set A, while in set C, the large outlier, 245, increased the mean. ■

### 5.2.2 Measures of Variation

Once you obtain a set of data, there are some basic questions we want to answer about the data: How common are the data values? Do we have a lot of repeated values? How much difference is there among data values?


Figure 5.17: Measures of variation help us see dispersion of the data values

This section will teach different ways to think about the data's variation or spread. Are the numbers spread out, or are they close together? Are there any outliers? We will consider two different measures of variation: range and standard deviation. Then we will learn about a third measure known as the interquartile range.

**Definition 5.2.6** The **range** of a data set is the difference between the highest data value and the lowest data value.

■ **Example 5.30** Find the range of this data set:

$$\{52, 36, 49, 23, 46, 77, 55, 48\}$$

**Solution:**

The lowest number is 23, and the most significant value is 77, so the range is $77 - 23 = 54$. ■

While the range is a fundamental way of calculating the distance or spread of the data values, it is not very useful if the data has an outlier. We need more powerful tools to understand the spread and variation of the data. One tool we can use is a measure called **standard deviation**.

**Definition 5.2.7 — The Standard Deviation**

The **standard deviation** is a number that measures how far all of the data values are from the mean.

The equation for the sample standard deviation is:

$$S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$$

The standard deviation is always greater than or equal to zero. The standard deviation is slight when the data are all concentrated close to the mean, exhibiting little variation or spread. The standard deviation is more prominent when the data values are spread out from the mean, showing more variation.

Although there is not a direct relationship between the range and standard deviation, there is a rule of thumb that can be useful to relate these two statistics. This relationship is sometimes referred to as the range rule for standard deviation.

**Definition 5.2.8 — Range Rule of Thumb**

$$S_x \approx \frac{\text{Range}}{4} = \frac{\text{Max-Min}}{4}$$

This quick estimate of the standard deviation is pretty accurate with most data sets (especially data sets close to a normal distribution).

**Definition 5.2.9 — The Variance**

**Variance** is a measure of dispersion, meaning it is a measure of how far a set of numbers is spread out from their average value or **mean**, and is denoted as $S^2$.

$$S^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

The reason to denote the variance as $S^2$ comes from the **standard variation** $S_x$ being the square root of the variance. A disadvantage of the variance for practical applications is that, unlike the standard deviation, its units will differ from the variable, which is why the standard deviation is more commonly reported as a measure of dispersion.

**■ Example 5.31 — Student Weight**

The following data set is the weight in pounds (lbs) of a group of 10 students. Find the sample variance and the sample standard deviation.

$$\{190, 300, 253, 271, 120, 176, 150, 200, 189, 202\}$$

**Solution**

First we find the mean;

$$\bar{x} = \frac{190+300+253+271+120+176+150+200+189+202}{10} = 205.1$$

Now we subtract each data value from the mean, square the answers, and add them together:

$$\begin{aligned}\sum(x-\bar{x})^2 &= (190-205.1)^2 + (300-205.1)^2 + (253-205.1)^2 + (271-205.1)^2 \\ &+ (120-205.1)^2 + (176-205.1)^2 + (150-205.1)^2 + (200-205.1)^2 \\ &+ (189-205.1)^2 + (202-205.1)^2 = 26,710.58\end{aligned}$$

To find the variance, we divide by 10-1 which is all data points minus one:

$$S^2 = \frac{26,710.58}{10-1} = 2967.84$$

To find the standard deviation, we take the square root of our variance:

$$S_x = \sqrt{2967.84} = 54.48$$

So the variance = 2967.84 and the standard deviation = 54.48

Now, let's estimate the standard deviation via the **range rule of thumb**.

First, the minimum and maximum values of the data set are **Min** $= 120$ and **Max** $= 300$, respectively. Hence the approximation is given by:

$$S_x \approx \frac{\text{Range}}{4} = \frac{\text{Max-Min}}{4} = \frac{180}{4} = 45$$

Notice the resulting number is not exact, but considering the data values, it is an excellent approximate measure of the spread of the data points. ■

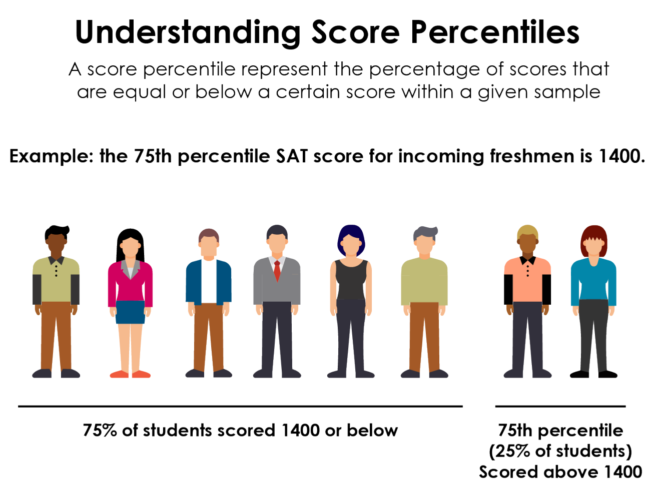
### 5.2.3 Percentiles and Quartiles

Although variance and standard deviation provide numerical measures of the dispersion of the data, it is sometimes difficult to grasp what they mean. We can get more detailed information using criteria called percentiles and quartiles.

**Definition 5.2.10** A **percentile** indicates a value below which a percentage of data falls when data are sorted into numerical order from smallest to largest. Percentages of data values are less than or equal to the $k^{th}$ percentile. For example, 15% of data values are less than or equal to the 15th percentile.

**Example 5.32 — PSAT scores**

When you take a standardized test, such as the PSAT, you are given both your score and the percentile of your score. For example, if it said you were in the $72^{nd}$ percentile, you did better than 72% of everyone else who also took the test. ■




**Example 5.33 — SAT scores in a school**

A school provides the following information about last year's incoming freshman class:

- The 75th percentile SAT scores for incoming freshmen was 1400

This tells us that 75% of the accepted students made 1400 or below on the test, and 25% of the accepted students scored above 1400. The $75th$ percentile means 25% of the accepted students made 1400 or above on the test. ■

In summary: Percentiles always correspond to lower data values and indicate percentages above and below a particular value.

**Definition 5.2.11 Finding the $k^{th}$ Percentile**
If n is the number of values in your data set and you want to find the data value in the $k^{th}$ percentile, you must first find the placement of that data point. This placement is also called the index and is denoted as i. To find the index $i$, we use:

- k = the percentile we are looking for
- n = the total number of data

Then the index formula is given by:

$$i = \frac{k}{100}(n+1)$$

The value of i is the placement of the number which is in the kth percentile if the values are written in ascending order.

**Example 5.34 — Academy Awards**

Listed below are 29 ages for Academy Award-winning best actors in order from youngest to oldest:

$18, 21, 22, 25, 26, 27, 29, 30, 31, 33, 36, 37, 41, 42, 47, 52, 55, 57, 58, 62, 64, 67, 69, 71, 72, 73, 74, 76, 77$

Figure 5.18: The Academy Awards

(a) Find the $70^{th}$ percentile.

(b) Find the $83^{rd}$ percentile.

**Solution:**

First, we solve (a), here:

- k = 70
- i = the index
- n = 29

$$i = \frac{k}{100}(n+1)$$
$$= \frac{70}{100}(29+1) = 21$$

The data value in the 21st position of the ordered data set is 64. The $70^{th}$ percentile is 64 years. Now, let's solve for (b), here:

- k = 83
- i = the index
- n = 29

$$i = \frac{k}{100}(n+1)$$
$$= \frac{83}{100}(29+1) = 24.9$$

The resulting index is not an integer, so we round it down to 24 and up to 25. The age in the 24th position is 71, and the age in the 25th position is 72. We find the mean of 71 and 72, and the $83^{rd}$ percentile is 71.5 years. ■

**Definition 5.2.12 — Quartiles** .

Quartiles are specific percentiles:

- The **first quartile, Q1**, is the $25^{th}$ percentile.
- The **third quartile, Q3**, is the $75^{th}$ percentile.
- The **median, M**, is the middle number, and it is the **second quartile, Q2**, and is the $50^{th}$ percentile.

To calculate quartiles and percentiles, we must order the data from smallest to largest. Quartiles divide ordered data into four equally sized sets of data. On the other hand, percentiles divide ordered data into hundredths. To score in the $90^{th}$ percentile of an exam does not mean, necessarily, that you received 90% on a test. It means that 90% of test scores are the same or less than your score, and 10% of the test scores are the same or greater than your test score.

Using the TI-83, 83+, 84, 84+ Calculator

To find the minimum, maximum, and quartiles:

Enter data into the list editor (Pres STAT 1:EDIT). If you need to clear the list, arrow up to the name L1, press CLEAR, and then arrow down.

Put the data values into the list L1.

Press STAT and arrow to CALC. Press 1:1-VarStats. Enter L1.

Press ENTER.

Use the down and up arrow keys to scroll.

Smallest value = 59.

Largest value = 77.

$Q_1$: First quartile = 64.5.

$Q_2$: Second quartile or median = 66.

$Q_3$: Third quartile = 70.

To construct the box plot:

Press 4:Plotsoff. Press ENTER.

Arrow down and then use the right arrow key to go to the fifth picture, which is the box plot. Press ENTER.

Arrow down to Xlist: Press 2nd 1 for L1

Arrow down to Freq: Press ALPHA. Press 1.

Press Zoom. Press 9: ZoomStat.

Press TRACE, and use the arrow keys to examine the box plot.

Figure 5.19: How to calculate quartiles with a TI calculator.

### ■ Example 5.35 — Quartile Computation

The following sample set corresponds to the GPAs of a group of students attending a college:

$$\{4.00, 1.76, 1.90, 2.12, 2.53, 1.39, 2.71, 3.00, 3.33, 3.71\}$$

Find the quartiles.

**Solution:**

First we must put the numbers in order:

$$\{1.39, 1.76, 1.90, 2.12, 2.53, 2.71, 3.00, 3.33, 3.71, 4.00\}$$

This data set has $n = 10$ observations. Since 10 is an even number, the median is the mean of the two middle data values:

$$\bar{x} = \frac{(2.53 + 2.71)}{2} = 2.62$$

Thus the second quartile is $Q_2 = 2.62$

Once we find the median, we can divide the sample set into two data sets, the **lower** half and the **upper** half. We can now compute each set's median to find the first and third percentiles, respectively.

Lower:

$$L = \{1.39, 1.76, 1.90, 2.12, 2.53\}$$

Upper:

$$U = \{2.71, 3.00, 3.33, 3.71, 4.00\}$$

Each set has an odd number of elements, so the median of each set is its middle observation. Thus, the first quartile is $Q_1 = 1.90$ (the median of L), and the third quartile is $Q_3 = 3.33$ (the median of U). ■

**Definition 5.2.13 — The Five-Number Summary**

In addition to the three quartiles, the two extreme values, the **minimum** value and the **maximum** value, are also useful in describing the entire data set. Together these five numbers are called the five-number summary of a data set.

$$X_{min}, Q_1, Q_2, Q_3, X_{max}$$

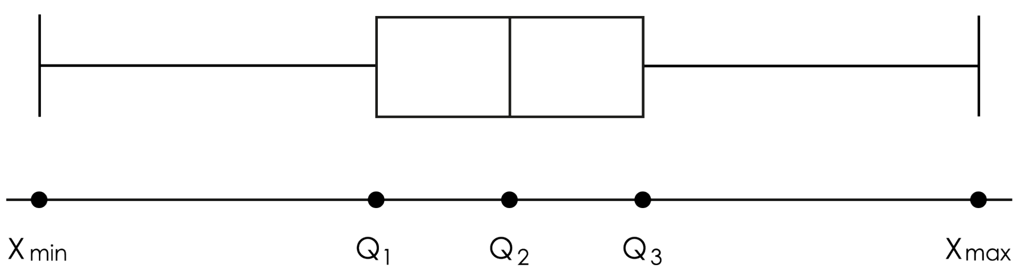



Figure 5.20: Box plot

The **five-number summary** is used to construct a plot also known as a **box-and-whisker plot**. Figure 5.20 shows such a graph. Each of the five numbers is represented by a vertical line segment. A box is formed using the line segments at Q1 and Q3 as its two vertical sides, then the median (or Q2) is the vertical line inside the box. Two horizontal line segments are extended from the vertical segments marking Q1 and Q3 to the adjacent extreme values. (The two horizontal line segments are called "whiskers.")

Regarding measures of variation, there is a third measure we have not yet discussed. This measure is the inter-quartile range.

**Definition 5.2.14 — Inter-Quartile Range**

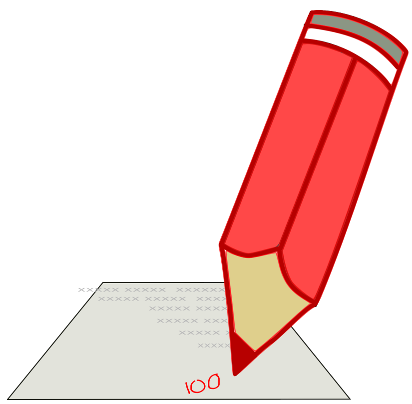
The **inter-quartile range**, or $IQR$, is the difference between the upper and the lower quartiles.

$$IQR = Q3 - Q1$$

The inter-quartile range tells us more than the range because we know the middle 50% of the data is between the quartiles. It gives us a better spread idea because it does not account for outliers.

■ **Example 5.36 — Test Scores**

Test scores for a college statistics class are compiled:




The morning class has the following scores:

- $\{99, 56, 78, 55.5, 32, 90, 80, 81, 56, 59, 45, 77, 84.5, 84, 70, 72, 68, 32, 79, 90\}$

The class held during the evening has the scores:

- $\{98, 78, 68, 83, 81, 89, 88, 76, 65, 45, 98, 90, 80, 84.5, 85, 79, 78, 98, 90, 79, 81, 25.5\}$

(a) Find the smallest and largest values, the median, and the first and third quartile for the morning class.

(b) Find the smallest and largest values, the median, and the first and third quartile for the evening class.

(c) For each data set, what percentage of the data is less than the first quartile? Between the first quartile and the median? Between the median and the third quartile? Greater than the third quartile? Greater than the first quartile?

(d) What is the interquartile range for each class?

**Solution:**

(a) First, we rearrange the data in ascending order and find the 5-number summary.

$$\{32, 32, 45, 55.5, 56, 56, 59, 68, 70, 72, 77, 78, 79, 80, 81, 84, 84.5, 90, 90, 99\}$$

- Min = 32
- Q1 = 56
- M = 74.5
- Q3 = 82.5
- Max = 99

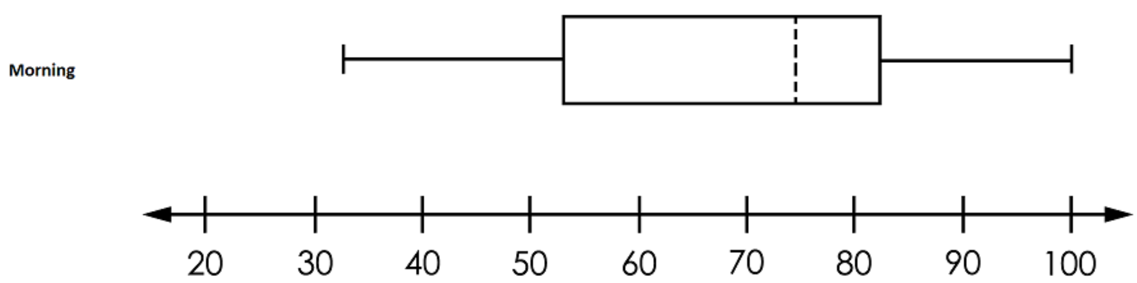



Figure 5.21: The box plot for the morning class.

(b). Again, we rearrange the data in ascending order and find the 5-number summary.

$$\{25.5, 45, 65, 68, 76, 78, 78, 79, 79, 80, 81, 81, 83, 84.5, 85, 88, 89, 90, 90, 98, 98, 98\}$$

- Min = 25.5
- Q1 = 78
- M = 81
- Q3 = 89
- Max = 98

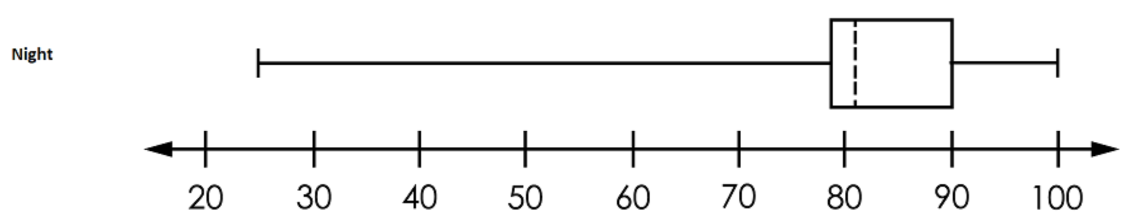



Figure 5.22: The box plot for the night class.

Let's look at the data again, but we will draw a line through the quartiles this time. If a quartile is between two numbers, draw the line between the numbers.

A: 32, 32, 45, 55.5, 56 | 56, 59, 68, 70, 72, | 77, 78, 79, 80, 81, | 84, 84.5, 90, 90, 99

B: 25.5, 45, 65, 68, 76, 78, 78, 79, 79, 80, 81, | 81, 83, 84.5, 85, 88, 89, 90, 90, 98, 98, 98

Notice that each section has 25% of the data.

(c) For each data set: 25% of the data falls below the first quartile, 25% of the data falls between the first quartile and the median, 25% falls between the median and the third quartile, and 25% falls above the third quartile. Because three of the sections are above the first quartile, 75% falls above Q1.

(d) For the morning class, the inter-quartile range is $Q3 - Q1 = 82.5 - 56 = 26.5$. For the night class the inter-quartile range is $Q3 - Q1 = 89 - 78 = 11$.

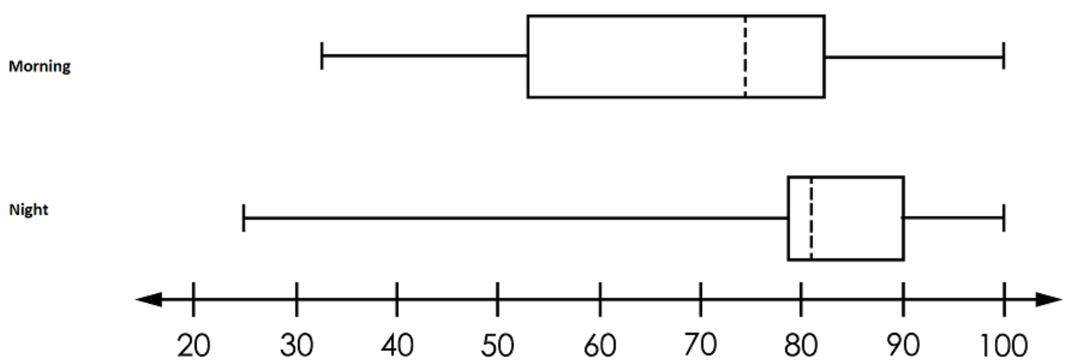



Figure 5.23: The box plots for the morning and night classes.

Notice the difference between their inter-quartile ranges. The middle 50%, or the values between Q1 and Q3, are much closer in the night class than in the morning class. This information is important if the data is used to decide if a class needs more instruction. ∎

### 5.2.4 Skewness and Normal Distribution

So far, we have considered several ways to measure the dispersion of a set of data. We also learned about box plots, our first visual tool to analyze how close data points are to each other. Next, we will define how the data is organized similarly to graphing a function.

**Definition 5.2.15 — Distribution**.

The **distribution** of a set of data is a description, graph, or function of the data versus frequency, and shows the approximate shape and dispersion of the data when graphed.

**Definition 5.2.16 — Skewness**.

**Skewness**, in statistics, is the degree of asymmetry observed in a distribution.

A distribution with the central location to the left and a tail off to the right is said to be **positively skewed** or **skewed to the right**. A positively skewed distribution means a majority of the observations are small relative to the rest of the distribution.

A distribution with the central location to the right and a tail off to the left is said to be **negatively skewed** or **skewed to the left**. A negatively skewed distribution means most of the observations are large relative to the rest of the distribution.

When there is an outlier, the mean is affected, but not the median or the mode. The mean, median, and mode are the same in a perfectly symmetric graph. In a **positively skewed** (or right-skewed) distribution, the mean is pulled up by the outlier in the positive (right) direction. In a **negatively skewed** (or left-skewed) distribution, the mean is pulled down by the outlier in the negative (left) direction.

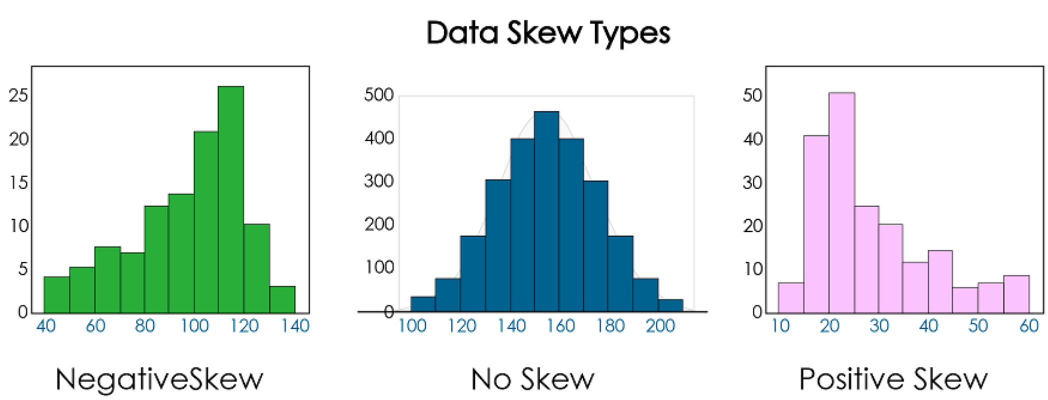



Figure 5.24: Skewness

In Figure 5.24, the mode is always the tallest bar. In the symmetric image, the median (middle value) and the mean are the same as the mode. In contrast, in the positively skewed image, the outlier pulls the mean to the right of the tallest bar. In the negatively skewed image, the outlier(s) pull the mean to the left of the most elevated bar.

**Example 5.37 — Influence of Outliers**

The following distributions (sets of data) have the same median: 42. Without calculating anything, which of these distributions will have the largest mean, and which will have the smallest mean?

- $A = \{34, 35, 37, 39, 41, 42, 43, 44, 46, 48, 49\}$
- $B = \{20, 34, 35, 37, 39, 41, 42, 43, 44, 46, 48\}$
- $C = \{34, 35, 37, 39, 41, 42, 43, 44, 46, 48, 80\}$

**Solution:**

The data in the three distributions lies between 34 and 49. In B, there is an outlier of 20. Also, C has an outlier of 80, so they have the smallest and highest means, respectively. ■

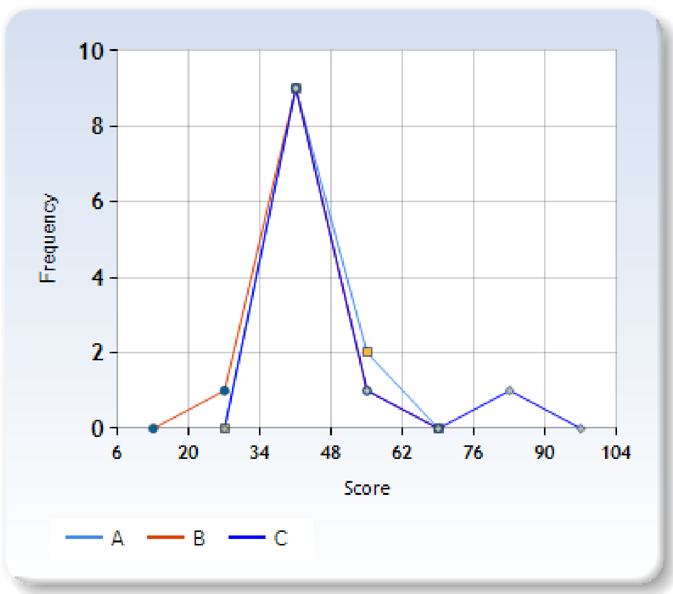



Figure 5.25: The data versus frequency plots for sets A, B and C

Notice that the distributions are not symmetric and have a slight skewness to the left (or positively skewed). The outliers in B and C will make the tails for these distributions look longer.

### 5.2.5 The Normal Curve

Often in nature, we find data values that, when graphed, take on a bell-shaped curve called a **normal curve**. A few simple ideas can help us better understand where the data lie in a normal distribution.

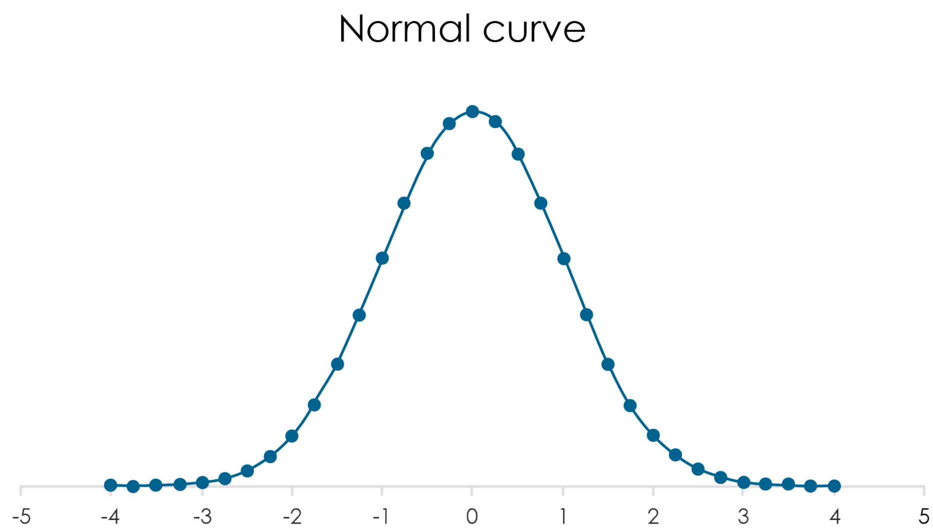



Figure 5.26: The normal curve

In figure 5.26 the highest point of the graph (the mode) is directly in the middle, so the graph is symmetric. This point is not only the mode but also the median and the mean; this is always true for a normal distribution.

The normal curve is a symmetric distribution about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graphical form, a normal distribution will appear as a bell-shaped curve.

- In a **normal distribution**, the mean is denoted $\mu$ and the standard deviation is denoted $\sigma$.
- In a **standard normal distribution**, is a specfic normal distribution in which the mean is zero and the standard deviation is 1.
- Normal distributions are symmetrical, but not all symmetrical distributions are normal.

Per symmetry of a normal distribution, the mean is equal to the median. 50% of the data is above, and 50% is below the mean. If we include one, two, or three standard deviations from this central point, we can know how much data will be covered by the curve thanks to the **empirical rule** or the $68\% - 95\% - 99.7\%$ law.

**Definition 5.2.17 — The Empirical Rule**

When data is normally distributed, the following can be expected:

1. Approximately 68% of the data lie within one standard deviation of the mean.
2. Approximately 95% of the data lie within two standard deviations of the mean.
3. Approximately 99.7% of the data lie within 3 standard deviations of the mean.

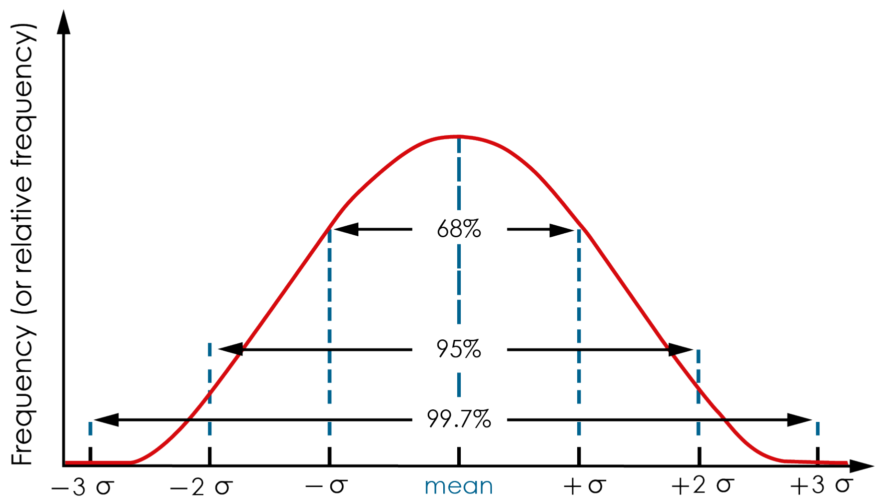



Figure 5.27: The empirical rule

Due to the empirical rule, we can also find percentages of data between different parts of the curve, for example:

1. The percentage between the mean value and the mean plus one standard deviation is 34%
2. The percentage between the mean value and the mean minus one standard deviation is 34%
3. The percentage between mean value minus two standard deviations and the mean minus one standard deviation is 13.5%
4. The percentage between the mean value minus three standard deviations and the mean minus two standard deviations is 2.1%
5. The percentage of data after $\mu + 2\sigma$ added to the before $\mu - 2\sigma$ is 4.2%

**■ Example 5.38 — Snowfall** .

In this example, the depth of snow in a yard is normally distributed, with $\mu = 2.5''$ and $\sigma = 0.25''$.

(a) What is the probability that a randomly chosen location will have a snow depth between 2.25 and 2.75 inches?

(b) What is the probability that a randomly chosen location will have a snow depth between 2 and 3.25 inches?

■

Figure 5.28: Snow in a yard

**Solution:**

(a) 2.25 inches is $\mu - \sigma$, and 2.75 inches is $\mu + \sigma$, so the area encompassed approximately represents $34\% + 34\% = 68\%$.

The probability that a randomly chosen location will have a depth between 2.25 and 2.75 inches is $68\%$.

(b) 2 inches is $\mu - 2\sigma$, and 3.25 inches is $\mu + 3\sigma$, so the area encompassed approximately represents $47.5\% + 49.85\% = 97.35\%$.

The probability that a randomly chosen location will have a depth between 2 and 3.25 inches is $97.35\%$.

### 5.2.6 Exercises 5.2

**Problem 5.8** A mom at a playground wants to know the average age of the children playing on the playground. Their ages are:

$$\{2,6,4,9,4,7,3,10,9,9,5,3,3,7,6,16\}$$

(a) What is the mean of the children's ages?

(b) What is the median of the children's ages?

(c) What is the mode of the children's ages?

(d) Which of the above statistics do you think will be most useful for the mom and why?

(e) What is the range of the ages? Approximate the standard deviation with the range rule.

(f) What is the standard deviation of the ages?

(g) Notice the difference in the mean and the median. Which is a larger number and why?

(h). Is the distribution of the data skewed? How do you know?

**Problem 5.9** A used car salesman wants to know how many cars he sells weekly. He tracks his sales for 15 weeks and finds that the amount he sold each week was as follows:

$$\{2, 11, 7, 3, 10, 5, 6, 6, 8, 4, 20, 2, 0, 4, 7\}$$


(a) Write the 5-number summary: minimum, Q1, median, Q3, maximum.

(b) Draw a boxplot of the data.

(c) To get a good idea of a normal week, the salesman wants to know the mean and standard deviation of the data.

(d) To get a good idea of a normal week, the salesman wanted to know the percentile of the 50%. The middle 50% falls between what two numbers?

(e) Since the salesman studied statistics in college, he knows that each section of the boxplot contains 25% of the data. Why, then, do the sections have different lengths?

**Problem 5.10** The same used car salesman decides to track another set of 15 weeks and finds:

$$\{12, 1, 7, 13, 0, 5, 4, 4, 4, 14, 20, 12, 10, 14, 7, 11\}$$

answer questions (a), (b), and (c) from the previous problem with this new data.

**Problem 5.11** A data set with no context is given below:

$$\{27, 43, 58, 92, 104, 119, 121, 123, 127, 140, 152, 155, 168, 190, 205, 212, 230\}$$

Find each of the following:

(a) Which number is the 23rd percentile?

(b) Which number is the 90th percentile?

(c) 121 is what percentile?

(d) What is another word for each of the following: the 25th percentile, the 50th percentile, and the 75th percentile?

(e) Can you make any interpretations about this data?

**Problem 5.12** Jack and James are the top players in a basketball league. The table below shows the number of points scored by Jack and James throughout 18 games in the league.

| Game | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Jack | 24 | 19 | 24 | 20 | 25 | 21 | 25 | 21 | 26 | 22 | 26 | 22 | 27 | 23 | 27 | 23 | 28 | 24 |
| James | 23 | 19 | 26 | 24 | 20 | 21 | 23 | 24 | 25 | 29 | 24 | 22 | 25 | 24 | 27 | 23 | 24 | 28 |

An award is to be given to the more consistent player of the two. We must make a decision mathematically.

(a) Design as many measures of consistency as you can to determine the more consistent player. Show all work.

(b) Did you use mean, median, and mode? If not, calculate them for each player.

(c) Did you use the standard deviation above? If not, calculate it by estimating using the rule of thumb and then using the formula for each player.

(d) Did you use the 5-number summary above? If not, calculate them and graph the box plot for each player.

(e) Can you chart the frequencies? Are they symmetric? Is the data normally distributed?

(f) Which player is more consistent after all your analyses? Justify your answer.

**Problem 5.13** A survey finds that the prices paid for five-year-old Kia Rio cars are normally distributed with a mean of $\$10,500$ and a standard deviation of $500. Consider a sample of 20,000 people who bought five-year-old Kia Rio cars and answer the following questions (**tip: use the empirical rule**):


(a) How many people paid between $\$10,000$ and $\$11,000$?

(b) How many people paid less than $\$10,500$?

(c) How many people paid less than $\$10,000$?

(d) How many people paid more than $\$11,500$?

(e) How many people paid between $\$10,000$ and $\$11,500$?

**Problem 5.14** The weights of a certain animal have a weighted average of 70 lbs with a standard deviation of 2.5 lbs. Assuming the weights follow a normal distribution find:

(a) What weight is 2 standard deviations below the mean?

(b) What weight is 1 standard deviation above the mean?

(c) How much does the middle 68% weigh?

## 5.3 Visualization of Data

What is the data telling you? When looking at numerical data, it can be challenging to see the big picture. Because of this, tables and graphs have been developed to help us better understand the data. This section will look at some of the most common tables and charts that help us visualize data.

### 5.3.1 Frequency Tables

A **frequency table** is a table with a column for the data values and another for how often each value occurs.

#### ■ Example 5.39 — Frequency in Grades

A teacher gave a test and the grades on the test were:

$$A,A,A,A,A,B,B,B,B,B,B,B,B,C,C,C,C,D,D,F,F,F,F,F,F,F.$$

Make a frequency table for the data.

**Solution:**

| **Grade** | A | B | C | D | F | Total |
|---|---|---|---|---|---|---|
| **Frequency** | 5 | 8 | 4 | 2 | 7 | 26 |

■

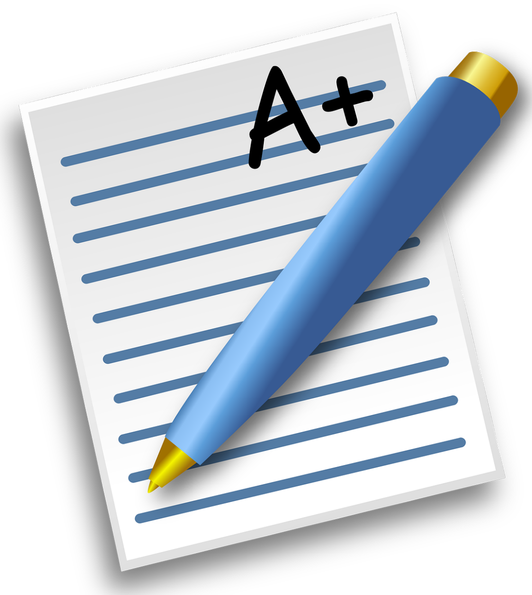
When numerical data has many possible values, we sometimes want to "bin" the data into intervals. Each of these bins is referred to as a **class**. In other words, specific numbers are put into groups together as categories.

**■ Example 5.40 — Binned Frequencies and Grades**

The same teacher from the previous example decided she needed to look more closely at the grades. She wanted to separate the low A's from the high A's, the low B's from the high B's, etc. The following are the numerical grades the students received:

$$92,98,94,99,92,80,85,89,83,88,88,80,81,72,79,75,74,64,60,53,40,42,44,44,44,42$$

Make a frequency table with the bins 95-99, 90-94, 85-89, and so on until all the data is in the table.




**Solution:**

| **Grade** | 95-99 | 90-94 | 85-89 | 80-84 | 75-79 | 70-74 |
|---|---|---|---|---|---|---|
| **Freq** | 2 | 3 | 4 | 4 | 2 | 2 |

| **Grade** | 65-69 | 60-64 | 55-59 | 50-54 | 45-49 | 40-44 | Total |
|---|---|---|---|---|---|---|---|
| **Freq** | 0 | 2 | 0 | 1 | 0 | 6 | 26 |

Table 5.1: This table shows binned data

■

There are some characteristics you should notice about the binned data in the previous example:

- Each class is the same size (95-99 is the same length as 80-84).
- No interval was left out of the middle, even if there was no data in that class.
- The total should be the same as the table in the example before it.

When looking at data, it is often more helpful to see what fraction or percentage of the data is in each class. This fraction or percentage is called the **relative frequency**. Sometimes we want to know how many data values are above a certain point. That number is called the **cumulative frequency**.

**Definition 5.3.1 — Relative and Cumulative Frequencies**

- **Relative frequency** is the fraction (or percent) of the data in each class of the frequency table. It is found by dividing the frequency by the total number of data points.
- **Cumulative frequency** is the total of the frequency of that class and every class before it.

To illustrate this, we will use the table from our original example of grades. We will add two more columns to the table to include the relative and cumulative frequency.

| Grade | Frequency | Relative Frequency | Cumulative Frequency |
|---|---|---|---|
| A | 5 | $\frac{5}{26}$ | 5 |
| B | 8 | $\frac{8}{26}$ | 13 |
| C | 4 | $\frac{4}{26}$ | 17 |
| D | 2 | $\frac{2}{26}$ | 19 |
| F | 7 | $\frac{7}{26}$ | 26 |
| Total | 26 | 1 | 26 |

Table 5.2: This table illustrates how to find relative and cumulative frequencies.

Notice that the column of fractions (or decimals) will always have a sum of 1. But how was the cumulative frequency column filled? Let's look at this more closely:

- In the row for grade A, the frequency is five, and there is nothing before it, so the cumulative frequency is also 5
- In the row for grade B, the frequency is eight, and we add any frequencies that come before it: 8+5=13

- In the row for grade C, the frequency is four, and we add any frequencies that come before it: 4+8+5=17
- In the row for grade D, we do the same and have 2+4+8+5=19
- In the row for grade F, we do it again and have 7+2+4+8+5=26, which is the total number of grades in the data set.

### 5.3.2 Graphs and Qualitative Data

There are two main graphs used for qualitative data: **bar graphs** and **pie graphs**. In the two previous examples, both sets of data were grades. However, one was letter grades, and one was numerical grades. Letter grades are a good example of qualitative data, so that we will use the data in table 5.2 as an example.

All statistical graphs should include a title or a caption. We should label the axis. In this example, the horizontal axis is the letter grades, and the vertical axis is the frequency of each letter grade. Figure 5.29 shows a bar graph of our data. Notice the frequencies are shown as the heights of the bars.

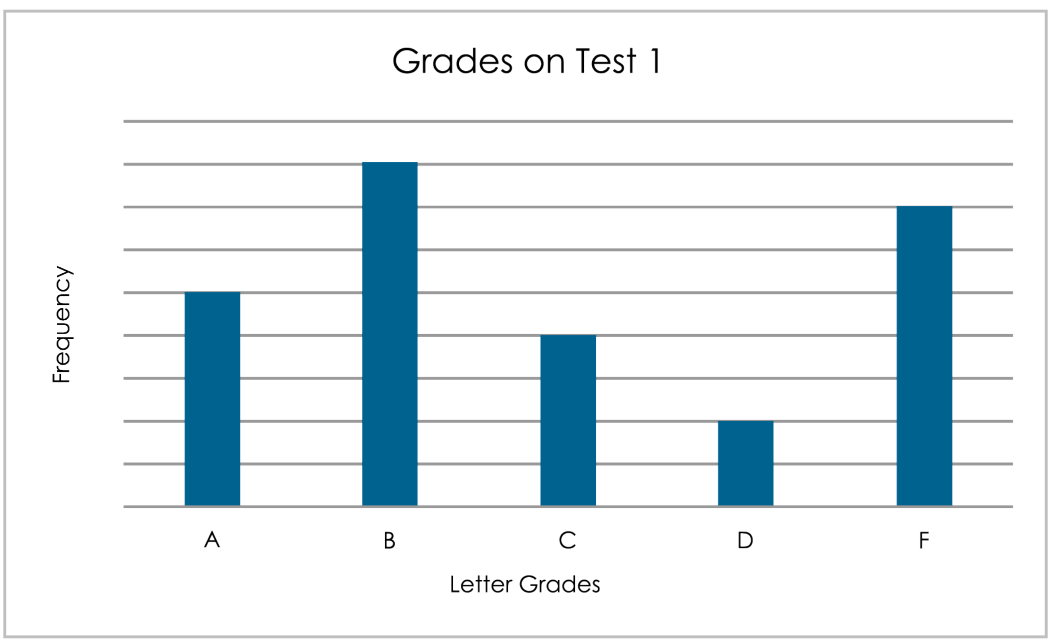



Figure 5.29: Bar graph of test grades

In some cases, we would instead compare the relative frequencies. This kind of frequency data is shown using a pie graph. In figure 5.30, the data are presented as percentages. Each of these percentages was found by converting the fractions from the table into shares. At the bottom of the graph, a key is included to explain what each piece of the pie represents. There should always be a title or caption with a chart.

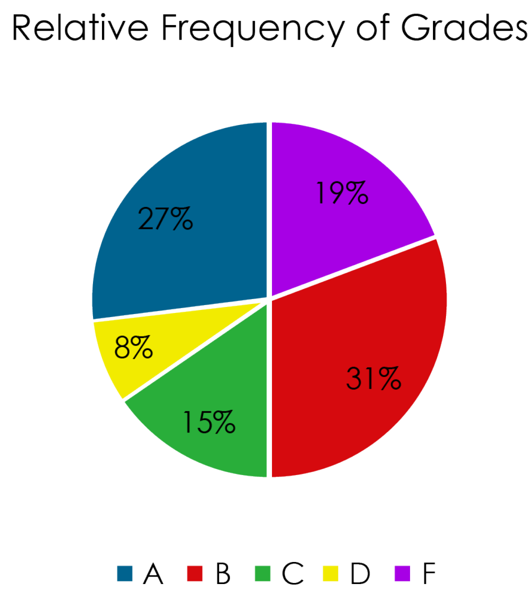



Figure 5.30: Pie graph of test grades

Notice the wedges of the pie chart have different sizes according to the percentage or frequency shown. The size is determined by equating 100% to 360°. Hence 50% would correspond to 180°, 30% would correspond to 102°, 25% would correspond to 90°, 10% would correspond to 36°, and so on. A proportion is used to find the number of degrees in the circle corresponding to the percentage.

Below is a table comparing the number of part-time and full-time students enrolled at Georgia Gwinnett College and Kennesaw State University for the fall 2020 semester.

| GGC | Number | Percent | KSU | Number | Percent |
|---|---|---|---|---|---|
| Full-time | 7620 | 65.5% | Full-time | 26,779 | 71% |
| Part-Time | 4007 | 34.5% | Part-time | 11,028 | 29% |
| Total | 11,627 | 100% | Total | 37,807 | 100% |

Table 5.3: GGC versus KSU 2020 enrollments

It is a good idea to look at various graphs to see which is the most helpful in displaying the data. We might choose what we think is the "best" graph depending on the data and the context. Our choice of the graph will depend on the information we are trying to convey, what better illustrates the features of the data, and what complements the narrative of our conclusions. Look at figures 5.31 and 5.32, and decide which one seems to better show the data. First, look at the pie chart:

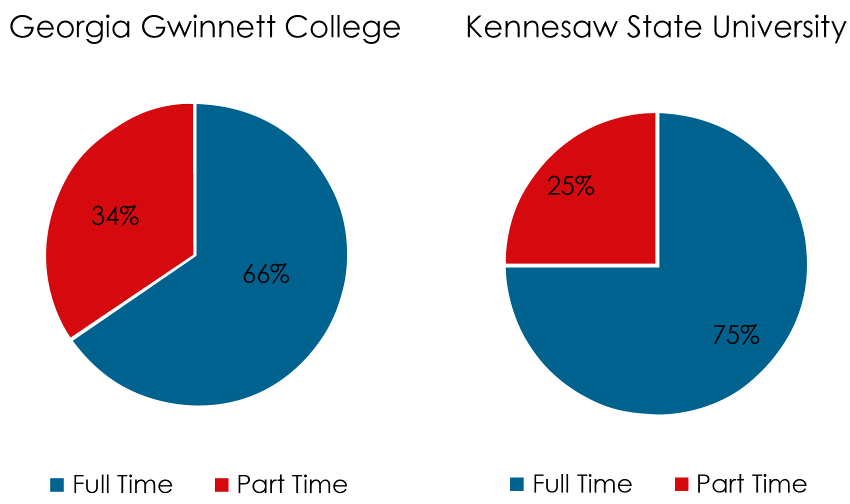



Figure 5.31: Public data from Georgia Gwinnett College and Kennesaw State University

We can see the same information in table 5.3 as bar graphs:

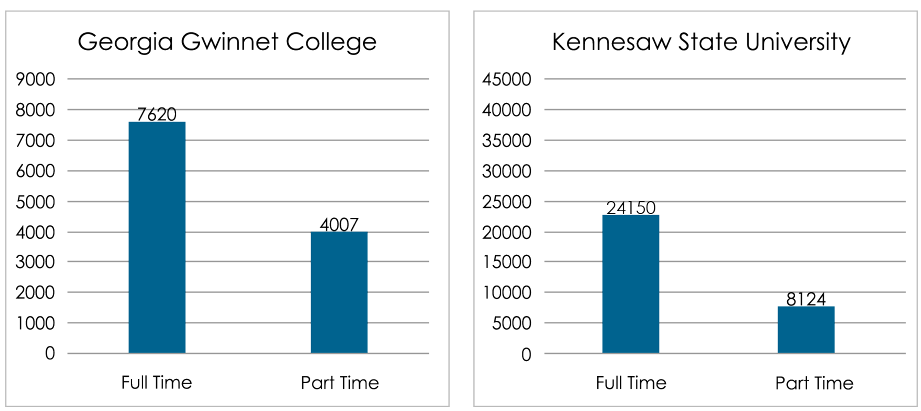



Figure 5.32: Enrollment data from GGC and KSU

A key characteristic of bar graphs is that the x-axis represents categories and does not include a numerical scale. Still, the y-axis is numerical and shows frequency counts.

There is a special kind of bar graph, called a **Pareto chart**, which can help compare the sizes of the bars. A Pareto chart is different because the bars are organized from largest to smallest. In figure 5.33 the students' ethnicities from Georgia Gwinnett College are shown. Notice how well this graph highlights the diversity of GGC.

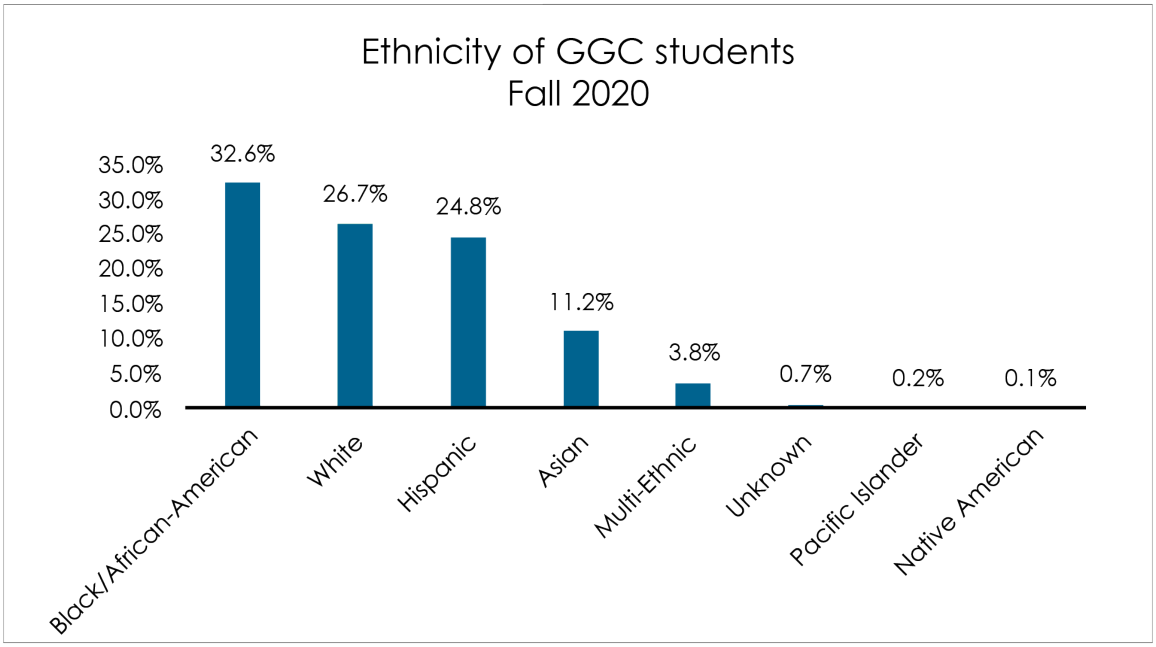



Figure 5.33: Data from the public website of Georgia Gwinnett College

### Pareto Chart Steps

The steps to build a **Pareto chart** are as follows:

1. Decide what categories you will use to group objects.
2. Decide what measurement is appropriate. Typical measurements are frequency, quantity, cost, and time.
3. Decide what time the Pareto chart will cover: One work cycle? One full day? A week?
4. Collect the data, recording the category each time, or assemble data that already exists.
5. Subtotal the measurements for each category.
6. Determine the appropriate scale for the measurements you have collected. The maximum value will be the largest subtotal from step 5. (If you do optional steps 8

and 9 below, the maximum value will be the sum of all subtotals from step 5.) Mark the scale on the left side of the chart.

7. Construct and label bars for each category. Place the tallest at the far left, then the next tallest to its right, and so on. If there are many categories with small measurements, we can group them as "other."

8. Calculate the percentage for each category: the subtotal divided by the total for all categories. Draw a right vertical axis and label it with percentages. Be sure the two scales match. For example, the left measurement corresponding to one-half should be precisely opposite 50% on the right scale.

9. Calculate and draw cumulative sums: add the subtotals for the first and second categories and place a dot above the second bar indicating that sum. Add the subtotal for the third category to that sum, and set a dot above the third bar for that new sum. Continue the process for all the bars. Connect the dots, starting at the top of the first bar. The last dot should reach 100% on the right scale.

#### Bar Graph Steps

The steps to build a **bar graph** are as follows:

1. Find the range in values. What units are used? What is the greatest value? What is the least value?

2. Determine a scale. Using graph paper, start with 1 box = 1 unit. What is the length (or height) of the longest bar? Will it fit? If not, change the scale and try again.

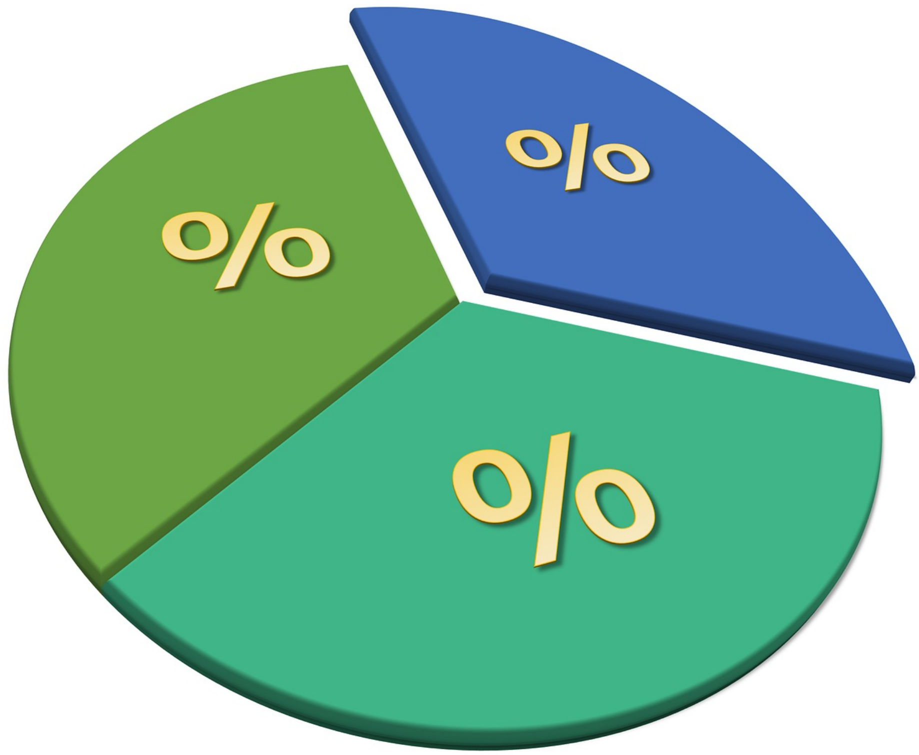
3. Label the graph. Mark each rectangle along the scale. Label the marks by the units they represent. Then decide how wide each bar should be. How much space will you allow between each bar?

4. Draw the bars. Mark where each bar starts and write the labels. Use your scale to determine the height of each bar.

5. Now, draw the bars on your graph. Check any two bars with values that are close in value. Will the bars show the difference?

6. Give the graph a title.

### Pie Chart Steps




The steps to build a **pie chart** are as follows:

1. First, put your data into a frequency table, then add up all the values to get a total.
2. Next, divide each value by the total and multiply by 100 to get a percent.
3. Now, you can figure out how many degrees for each "pie slice" (commonly called a sector). A full circle has 360 degrees, and we use this to calculate the degrees in each sector concerning the percentage.
4. Draw a circle.
5. Use a measuring tool, such as a compass, to mark the degrees of each sector.
6. Create labels for each of the sectors.

### 5.3.3 Visualizing Quantitative Data

In this section, we will discuss four graphs to visualize quantitative data. The boxplot (or box and whisker plot) was discussed earlier in this chapter. We will now consider histograms, line graphs, and stem-and-leaf plots. We will do all of these using the data from figure 5.1.

A **histogram** looks like a bar graph, but here are some important differences. The most obvious difference is that in a histogram, the bars touch, while in a bar graph, the bars are spaced apart. To create a histogram, we must first break the data into classes, just as we did in figure 5.1. The horizontal axis should be a number line, showing the range of numbers used to bin the data. In this case, our numbers must at least include 40 to 99. In figure 5.34, the data is presented as a histogram.

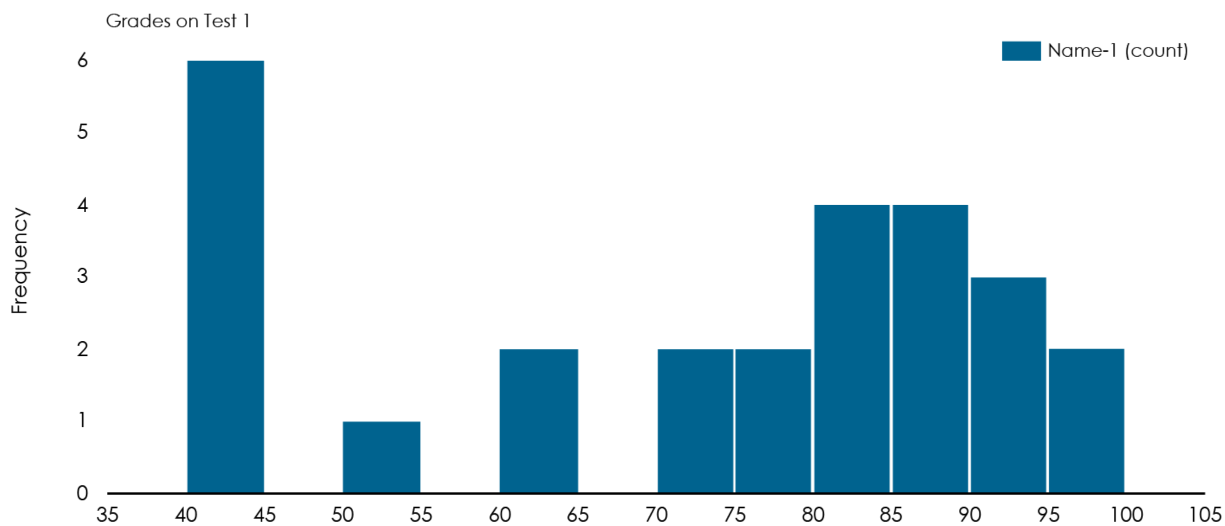



Figure 5.34: Tools like : `https://www.statskingdom.com/histogram-maker.html` can help you create histograms online.

#### Histogram Steps

Before we draw our histogram, we need to obtain basic summary statistics from our data set.

1. Create the **bins** or **classes**.

   First, find the maximum and minimum data values in the set. We can compute the range from these numbers by subtracting the minimum value from the maximum value. Next, use the range to determine the width of our classes. There is no set rule, but as a rough guide, the range should be divided by five for small data sets and 20 for larger sets. These numbers will give a class width or bin width. We may need to round this number or use some common sense.

2. Once the class width is determined, we choose a class that will include the minimum data value. We then use our class width to produce subsequent classes, stopping when creating a class that contains the maximum data value.

3. Frequency tables:
   Once we have determined our classes, the next step is making a frequency table. Begin with a column that lists the classes in increasing order. The next column should have a tally for each of the classes. The third column is for each class's data count or frequency. The final column is for the relative frequency of each class; this indicates what proportion of the data is in each class.

4. The drawing process:

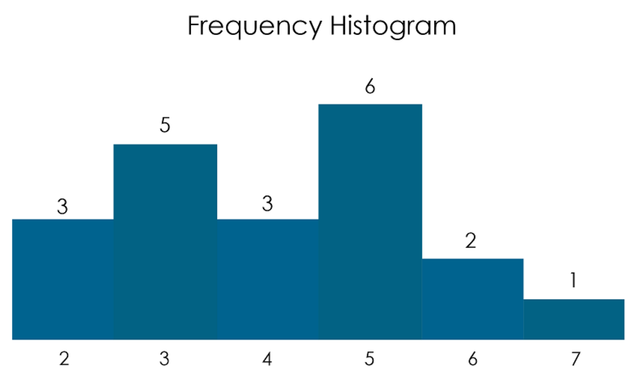



Figure 5.35: Histograms should have a clearly defined axis and scale. If it looks like this, it is not complete.

- Draw a horizontal line. This will be where we denote our classes.
- Place evenly spaced marks along this line that correspond to the classes.
- Label the marks so that the scale is clear and give a name to the horizontal axis.
- Draw a vertical line just to the left of the lowest class.
- Choose a scale for the vertical axis that accommodates the class with the highest frequency.
- Label the marks so the scale is clear, and label the vertical axis.
- Construct bars for each class. The height of each bar should correspond to the frequency or relative frequency of the class at the base of the bar.

Another graph that uses the data in a similar way is a **line graph**. The difference is that in a line graph, the midpoint of the top of each bar from a histogram is plotted as a point, and the points are connected with straight lines. Figure 5.36 shows a line graph displaying the same data as the histogram in Figure 5.34. We would typically use a histogram for this data because line graphs are most commonly used to show change over time.

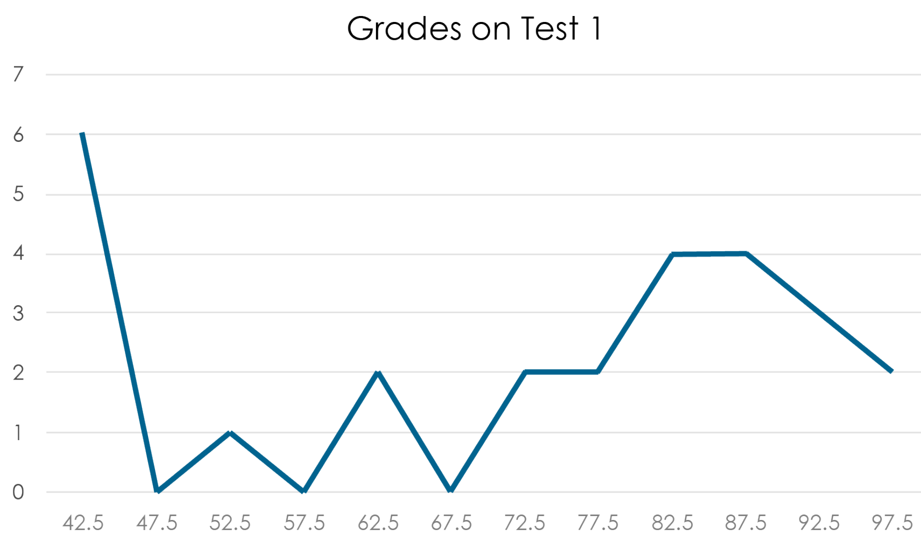



Figure 5.36: This line graph matches the histogram in figure 5.34. Although the figures are similar, the line graph emphasizes the change from grade bin to the next grade bin.

The steps to creat a **line graph** are as follows:

1. First, define your quantitative variables, x and y.
2. Next, define the x-axis with the appropriate units. Typically, line graphs represent change with respect time.
3. Then define the scales for both variables, x and y.
4. Plot the points corresponding to the data and each of the $(x, y)$ coordinate points.
5. Draw lines to connect the neighboring points.
6. Create labels for each of the axes.

It's important to note that line graphs are appropriate only when both the x and y axes display ordered (rather than qualitative) variables. Although histograms can be used in this situation, line graphs are better at comparing changes over time.

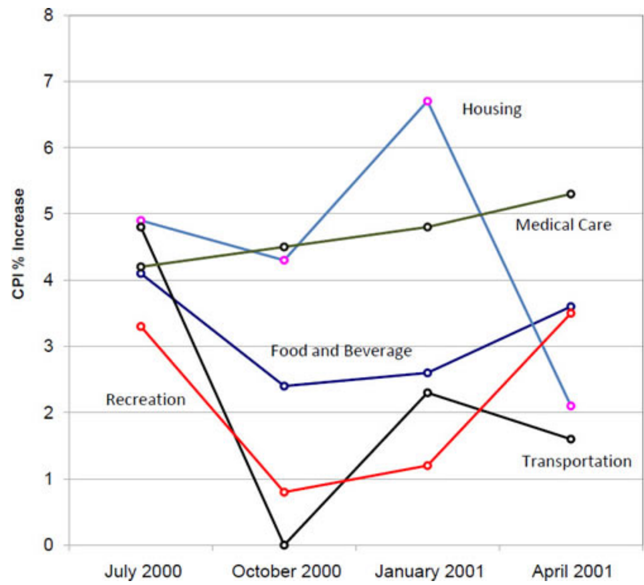



Figure 5.37: A line graph of the percent change in five components of the CPI over time.

Figure 5.37 shows the percentage increases and decreases in five components of the Consumer Price Index (CPI). The graph makes it easy to see that medical costs had a steadier progression than the other components.

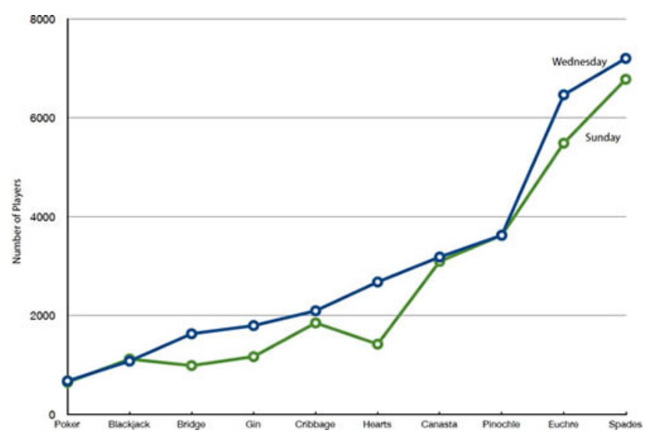



Figure 5.38: A line graph, inappropriately used, depicting the number of people playing different card games on Sunday and Wednesday.

It is inappropriate to use a line graph when the x-axis contains qualitative variables. Figure 5.38 inappropriately shows a line graph of card game data from Yahoo, discussed in the section on qualitative variables. The problem with Figure 5.38 is it gives a false impression that the games are naturally ordered numerically.

Another important graph that we introduced in section 5.2 is the **box plot**. We need the five-number summary of data introduced in Section 5.2 to draw this graph. The lines of the box are drawn at the 1st, and 3rd quartiles, the line inside the box is the median, and the whiskers go out to the lowest and highest values in the data set.

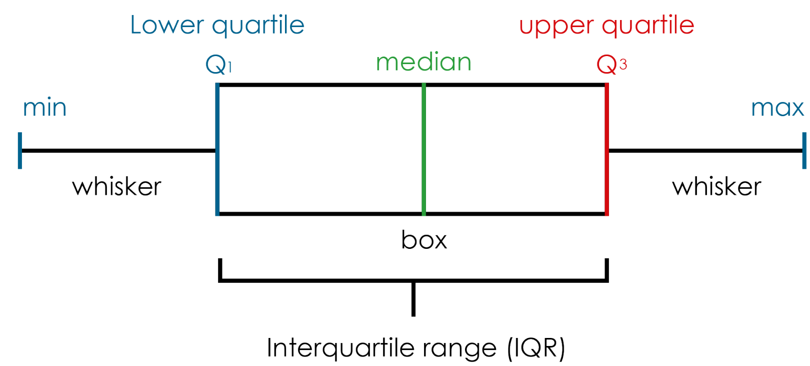



Figure 5.39: A box plot.

Figure 5.40 shows a box plot of the same data we have used for many examples in this chapter (Grades of Test 1). The box plot can inform the teacher of the middle 50% of her students who had grades between 53 and 88 (Q1 and Q3).

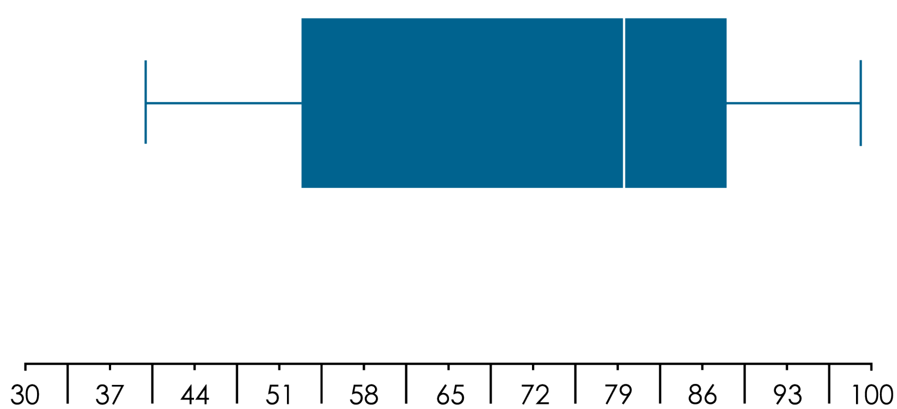



Figure 5.40: `http://www.shodor.org/interactivate/activities/BoxPlot/` was used to create this box plot.

Next, we introduce the **stem-and-leaf plot**. In this kind of graph, the "stem," or the numbers on the left side of the line, is the first digit of the data. The "leaf," or the numbers on the right side of the line, is the second digit of the data. For example, if all of our data between 80 and 89 are the scores 82, 83, and 85, then there would be one row of the stem-and-leaf plot that has an 8 to the left of the line, and the 2 and 3 and 5 (representing all of the data in the 80's) on the right.

In figure 5.41, we created a stem-and-leaf graph for the same data we used for most examples in this chapter (Grades of Test 1).

$$92, 98, 94, 99, 92, 80, 85, 89, 83, 88, 88, 80, 81, 72, 79, 75, 74, 64, 60, 53, 40, 42, 44, 44, 44, 42$$

| Stem | Leaf | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 4 | 0 | 2 | 2 | 4 | 4 | 4 | | |
| 5 | 3 | | | | | | | |
| 6 | 0 | 4 | | | | | | |
| 7 | 2 | 4 | 5 | 9 | | | | |
| 8 | 0 | 0 | 1 | 3 | 5 | 8 | 8 | 9 |
| 9 | 2 | 2 | 4 | 8 | 9 | | | |

Figure 5.41: `https://www.statology.org/stem-and-leaf-plot-generator/` was used to create the stem-and-leaf plot.

There are a few important characteristics to note about a stem-and-leaf plot:

- The numbers are in order. The top line in figure 5.41 is read as 40, 42, 42, 44, 44, 44. The ordered numbers make it easy to find the median (and the quartiles, too). There are 26 numbers, so the median is between the 13th and 14th values. The 13th value is 79, and the 14th value is 80. So, the median is $\frac{79+80}{2} = 79.5$.

- If you turn the stem-and-leaf plot on its side with the stem side down, it resembles a histogram. Try it with figures 5.41 and 5.34.

- We can use a stem-and-leaf plot to organize your data; this makes it easier to decide which kind of graph would be most appropriate for the situation.

You can also make a stem-and-leaf plot for two data sets to compare them. In this case, the stem is in the middle with the leaves for each data collection on either side.

Let's assume the same teacher has two classes who took Test 1, and she wants to compare the classes. Here are the grades for the second class:

$$80, 82, 87, 97, 99, 86, 88, 88, 88, 93, 94, 94, 73, 76, 62, 74, 70, 70, 72, 70, 69, 77$$

Figure 5.42shows the two classes compared. The original class is on the right, and the second class is on the left.

| | | |
|---:|:---:|:---|
| 9 2 | 4 | 0 2 2 4 4 4 |
| 7 6 4 3 2 0 0 0 | 5 | 3 |
| 8 8 8 7 6 2 0 | 6 | 0 4 |
| 7 4 4 3 | 7 | 2 4 5 9 |
| | 8 | 0 0 1 3 5 8 8 9 |
| | 9 | 2 2 4 8 9 |

Figure 5.42: Stem-and-Leaf plot created with Microsoft Excel

Notice that while both classes have about the same median (79.5 and 80), one class did significantly better than the other.

In general, the steps to build a **stem-and-leaf plot** are summarized as follows:

1. Determine the smallest and largest numbers in the data.

2. Identify the stems.

3. Draw a vertical line and list the stem numbers to the left of the line.

4. Fill in the leaves to the right of the line.

### 5.3.4 Linear Regression and Correlation

Sometimes we want to know how or if two or more numerical variables are related. If there is a relationship, what is the relationship, and how strong is it? For example, is there a relationship between the grade on a math exam and the amount of time spent doing homework? Another example is the relationship between your income and your level of education or your years of experience.

One method of exploring these relationships involves "linear regression," with one independent variable, x, and a dependent variable, y. This technique involves looking at a data graph and how it centers around a line drawn on a coordinate plane. Another part of this method concerns "correlation," which measures how strong the relationship is between the two variables.

#### Cause and Effect

First, we need to understand when two variables are related; we can look for a function-based relationship between the independent and dependent variables.

- Independent variables are variables whose effects are measured (cause). Independent variables are not affected by any other variables.

- Dependent variables measure the effect of the independent variable (effect). These variables depend on what is happening with other variables.

Examples of dependent and independent variables:

| Examples of independent and Dependent Variables in Experimental Research | |
|---|---|
| **Independent Variables.** | **Dependent Variables.** |
| Type of treatment: different types of drug treatment or psychological treatment. | Behavioral Variables: Measure of adjustment, activity levels, eating behavior, smoking behavior. |
| Treatment Factors: brief vs long term treatment, inpatient vs outpatient treatment. | Physiological variables: the measure of physiological responses such as heart rate, blood pressure, and brain wave activity. |
| Experimental manipulations: types of beverages consumed (alcoholic vs nonalcoholic) | Self-report variables: the measure of anxiety, mood, or marital or life satisfaction. |

Figure 5.43: Dependent versus independent variables

We use a scatter plot graph to show the direction of a relationship between the variables. A clear direction occurs when either:

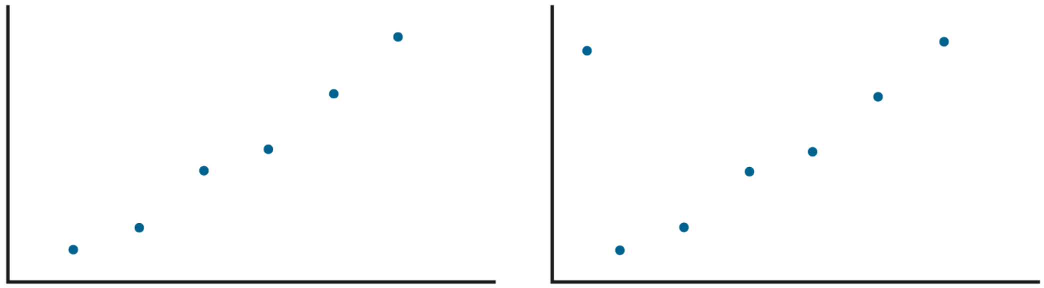
- Large values of one variable occur with large values of the other variable, or small values of one variable occur with small values of the other variable.

- Large values of one variable occur with small values of the other variable.

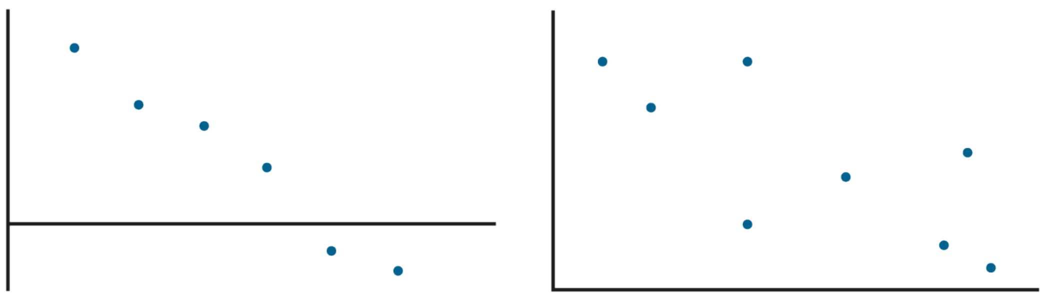
We can determine the **strength** of the relationship by looking at the scatter plot and seeing how close the points are to a line, a power function, an exponential function, or some other type of function. For a linear relationship, there is an exception where all the points fall on a horizontal line, meaning there is no relationship between the variables.

Linear patterns are common. The linear relationship is strong if the points are close to a straight line (except in the case of a horizontal line). If we think the points show a linear relationship, we draw a line on the scatter plot. We can calculate this line through a process called **linear regression**. However, we only calculate a regression line if one of the variables helps explain or predict the other variable. If x is the independent variable and y is the dependent variable, then we can use a regression line to predict y for a given value of x.

### Types of correlation



(a) Positive linear pattern (strong) (b) Linear pattern w/ deviaion



(a) Negative linear pattern (strong) (b) Negative linear pattern (weak)

Figure 5.44: Scatter plots

### Regression Equation

Data rarely fit a straight line exactly. Instead, we obtain rough predictions. Typically, a data set has a scatter plot that appears to "fit" a straight line. This line is called a **line of best fit or least-squares line**.

Before we learn how to find regression lines, let's review some essential characteristics of linear functions.

- The equation for a linear function is $y = mx + b$ (where $m$ is used for the slope and $b$ is the $y$-axis intercept)
- $y$ is the dependent variable
- $x$ is the independent variable

The slope of the line, m, describes how changes in the variables are related. It is essential to interpret the slope of the line in the context of the situation represented by the data.

### Interpretation of the slope

The slope of the best-fit line tells us how the dependent variable **y** changes for every one-unit increase in the independent **x** variable, on average.

Using the TI-83, 83+, 84, 84+ Calculator

To create a scatter plot:

1. Enter your X data into list L1 and your Y data into list L2.
2. Press 2nd STATPLOT ENTER to use Plot 1. On the input screen for PLOT 1, highlight On and press ENTER. (Make sure the other plots are OFF.)
3. For TYPE: highlight the very first icon, which is the scatter plot, and press ENTER.
4. For Xlist:, enter L1 ENTER and for Ylist: L2 ENTER.
5. For Mark: it does not matter which symbol you highlight, but the square is the easiest to see. Press ENTER.
6. Make sure there are no other equations that could be plotted. Press Y = and clear any equations out.
7. Press the ZOOM key and then the number 9 (for menu item "ZoomStat") ; the calculator will fit the window to the data. You can press WINDOW to see the scaling of the axes.

Figure 5.45: How to use calculator to draw scatter plots and obtain the least-squares equation.

### ■ Example 5.41 — Word processing costs

Aaron's Word Processing Service (AWPS) charges the following rate for services: \$32 per hour plus a \$31.50 one-time charge. The total cost to a customer depends on the number of hours it takes to complete the job. Find the equation that expresses the total cost in terms of the required hours. ■

**Solution:**

- Let x = the number of hours it takes to get the job done
- Let y = the total cost to the customer
- The \$31.50 is a fixed cost. If it takes $x$ hours to complete the job, then $(32)(x)$ is the cost of the word processing
- The total cost is: $y = 31.50 + 32x$

**Correlation does not always equal Causation**

Correlation is used to measure the linear relationship (strength) between two variables, whereas regression expresses the relationship as an equation.

For example, in the image below, the data indicate that ice cream sales increase with sunburn rates in summer. These variables are not related, and both are consequences of other phenomena.

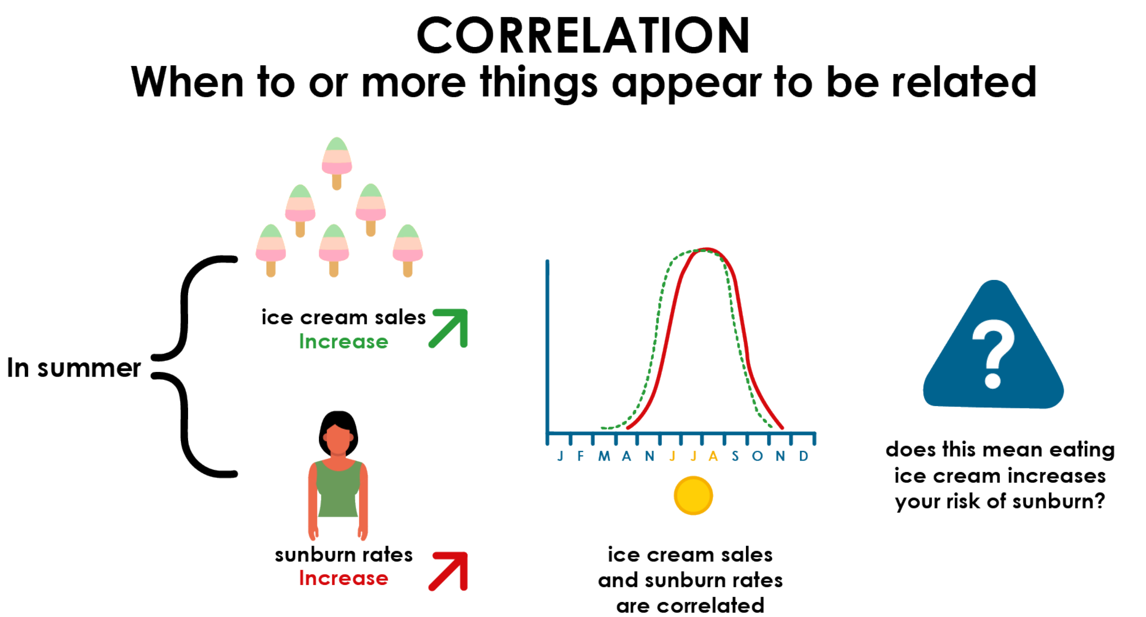



Figure 5.46: Correlation does not equal causation

A common misconception is that if there is a correlation between variables, one must have caused the other, but this is often not true. Let's look at an example where we expect a causal connection between two measurements. Consider buying a snake plant for your home. Snake plants are supposed to be easy to take care of because you can mostly ignore them.

Like most plants, snake plants need some water to stay alive. However, they also need just the right amount of water. Imagine an experiment where 1000 snake plants were grown in a house. Each snake plant is given a different amount of water per day, from zero teaspoons of water per day to 1000 teaspoons per day. We will assume that water is part of the causal process that allows snake plants to grow. The amount of water given to each snake plant per day can also be one of our measures. Imagine further that the experimenter measures snake plant growth every week, which is the second measurement. Imagine a scatter plot of weekly snake plant growth by tablespoons of water looks like.

#### Even when there is Causation, there might not be an obvious correlation

The first plant given no water at all would eventually die. It should have the least amount of weekly growth. How about the plants are delivered only a few teaspoons of water per day? If you imagine a scatter plot, with each dot being a snake plant's amount of water and its resulting growth, then you should visualize some dots starting in the bottom left-hand corner (no water and no plant growth) and moving up and to the right (a bit of water and a bit of growth). This could be just enough water to keep the plants alive so they will grow a little but not a lot.

Figure 5.47: Causation causes correlation but is not always easy to see.

As we look at snake plants getting more water, we should see more plant growth, right? Sure, but only up to a point. There should be a positive correlation with increasing plant growth as water per day increases. But what happens when you give snake plants too much water? Typically, they will die. At some point, the dots in the scatter plot will start moving back down again. Snake plants that get too much water will not grow very well.

We can conclude that **Correlation** describes an association between variables: when one variable changes, so do the other. A correlation is a statistical indicator of the relationship between variables that change together. That is, they co-vary. But this co-variation is not necessarily due to a direct or indirect causal link. **Causation** means that changes in one variable bring about changes in the other; there is a cause-and-effect relationship between variables. The two variables are correlated with each other, and there is also a causal link between them. **A correlation doesn't imply Causation, but Causation always implies Correlation**.

#### Confounding variable, or third variable problem

Anybody can correlate any two things that can be quantified and measured. For example, we could find a hundred people and ask them all sorts of questions like:

- How happy are you?
- How old are you?
- How tall are you?
- How much money do you make per year?
- How long are your eyelashes?
- How many books have you read?
- How loud is your inner voice?

Let's say we found a positive correlation between yearly salary and happiness. We could have just as easily computed the Correlation between satisfaction and annual salary. Would you infer that a yearly salary causes happiness if we found a correlation? Perhaps it does play a small part. But, something like happiness probably has a lot of contributing causes. Money could directly cause some people to be happy. But, more likely, money buys people access to all sorts of things, and some of those things might contribute to happiness. These "other" things are called third variables.

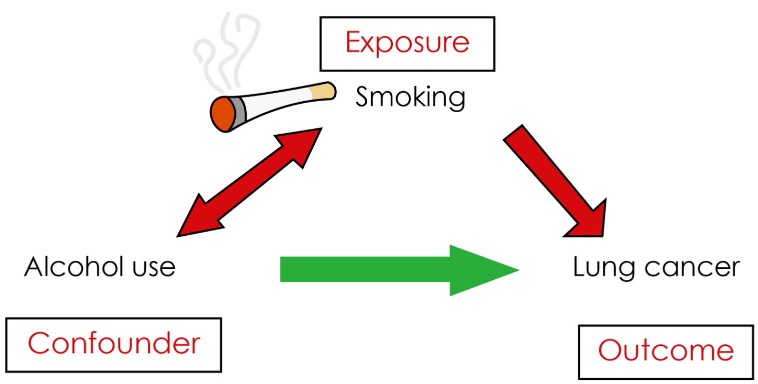



Figure 5.48: Drinking and smoking go hand in hand. However, lung cancer is not associated with alcohol use, although between patients, you can find indications of drinking. Alcohol use becomes a confounding factor.

For example, perhaps people living in more pleasant places in more expensive houses are happier than people in worse areas in cheaper homes. In this scenario, money is not causing happiness; it's the places and dwellings money buys. But, even if this were true, people can still be more or less happy in many different situations. The lesson is that a correlation can occur between two measures because a third variable is not directly measured. So, just because we find a correlation does not mean we can conclude anything about a causal connection between two measurements.

### 5.3.5 Exercises 5.3

**Problem 5.15** The lengths of 20 new born babies (in centimeters) were recorded as:

$$46.1, 55.9, 48.3, 52.4, 52.0, 50.7, 51.3, 48.9, 48.2, 51.3,$$
$$55.1, 55.6, 46.4, 53.4, 52.4, 51.7, 50.3, 52.0, 47.0, 49.8$$

Use this data to fill in the frequency table below.

| Length | Frequency | Relative Frequency | Cumulative Frequency |
|---|---|---|---|
| 46.0-47.9 | | | |
| 48.0-49.9 | | | |
| 50.0-51.9 | | | |
| 52.0-53.9 | | | |
| 54.0-55.9 | | | |
| total | | | |

**Problem 5.16** The following grades were from a midterm exam:

$$A, A, B, B, B, B, B, C, C, C, C, C, C, C, C, C, D, D, F, F, F$$

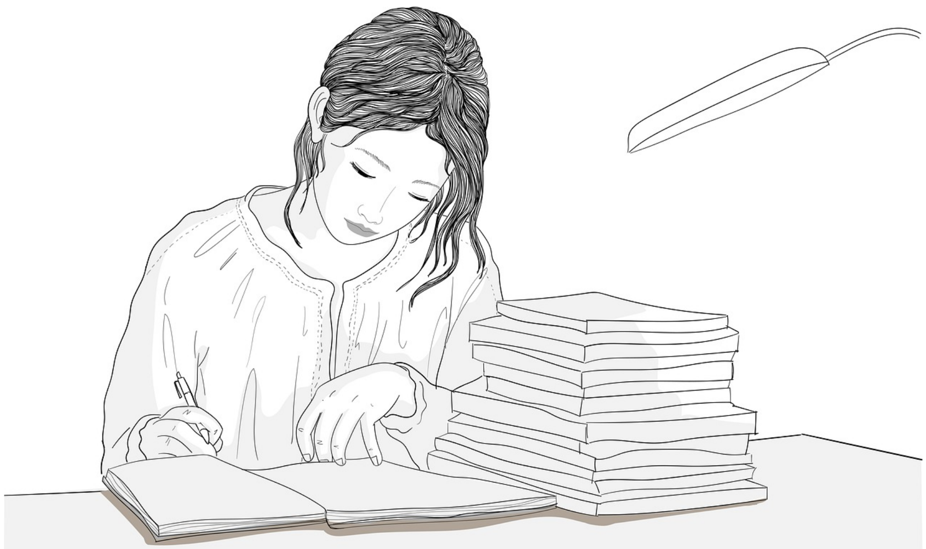
(a) Make a frequency table for this data. Make sure to include relative frequency and cumulative frequency.

(b) Draw a bar graph of this data.

(c) Draw a pie graph of this data.

**Problem 5.17** The following grades were from a midterm exam:

$$98, 95, 88, 86, 85, 84, 84, 79, 77, 77, 77, 73, 72, 72, 71, 71, 64, 62, 50, 47, 12$$



(a) Make a stem-and-leaf plot of this data..

(b) Make a frequency table for this data. Make sure to include relative frequency and cumulative frequency. The classes should be 90-99, 80-89, 70-79, 60-69, 50-59, 40-49, 30-39, 20-29, 10-19

(c) Make a histogram of this data.

(d) Make a line graph of this data.

(e) Make a box plot of this data.

**Problem 5.18** 67 GGC students were surveyed and asked what their place to purchase food on campus was. Their answers are in the frequency table below. Fill in the rest of the table.

| Location | Frequency | Relative Frequency | Cumulative Frequency |
|---|---|---|---|
| Dining Hall | 15 | | |
| GGC Cafe | 13 | | |
| Chik-Fil-A | 25 | | |
| Starbucks | 10 | | |
| The General Store | 4 | | |
| Total | 67 | | |

**Problem 5.19** Consider the following data

| | | | | |
|---|---|---|---|---|
| 9 | 92 | 68 | 77 | 80 |
| 70 | 85 | 88 | 85 | 86 |
| 93 | 75 | 76 | 82 | 100 |
| 53 | 70 | 70 | 82 | 85 |

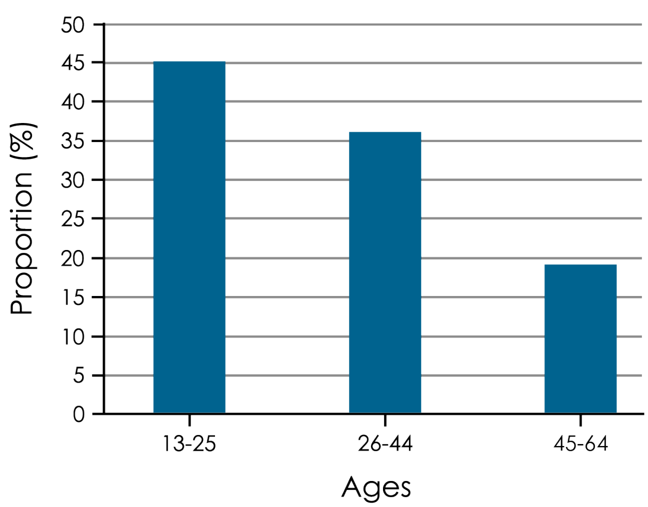
(a) Make a stem-and-leaf plot of this data..

(b) Make a frequency table for this data. Make sure to include relative frequency and cumulative frequency.

(c) Make a histogram of this data.

(d) Make a box plot of this data.

**Problem 5.20** Consider the following histogram:




(a) Describe a possible sample that would produce this histogram.

(b) Make a frequency table for the data.

(c) Calculate the mean and the five-number summary.

(d) What can you tell from this histogram about this data?

**Problem 5.21** Consider the following data:

| | | | | |
|---|---|---|---|---|
| 0 | 15 | 14 | 14 | 18 |
| 15 | 17 | 16 | 16 | 18 |
| 15 | 19 | 12 | 13 | 9 |
| 19 | 15 | 15 | 16 | 15 |

(a) Make a stem-and-leaf plot of this data.

(b) Make a frequency table for this data.

(c) Make a histogram of this data.

(d) Make a line graph of this data.

**Problem 5.22** Consider the following stem and leaf plot:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 5 | 3 | | | | | | |
| 6 | 8 | 9 | | | | | |
| 7 | 0 | 0 | 0 | 5 | 6 | 7 | |
| 8 | 0 | 2 | 3 | 5 | 5 | 5 | 8 |
| 9 | 2 | 3 | 6 | | | | |
| 10 | 0 | | | | | | |

(a) Describe the data sample that would produce this plot.

(b) Make a frequency table and a histogram.

(c) Calculate the mean and the five-number summary.

(d) Graph the box plot of the data.

**Problem 5.23** The price of a single issue of stock can fluctuate throughout the day. A linear equation representing the price of the stock for Shipment Express is

$$y = 15 - 1.5x$$

, where $x$ is the number of hours passed in an eight-hour trading day.

(a) What are the slope and y-intercept? Interpret their meaning.

(b) If you owned this stock, would you want a positive or negative slope? Why?

**Problem 5.24** A random sample of ten professional athletes produced the following data where x is the number of endorsements the player has, and y is the amount of money made (in millions of dollars).

| x | y | x | y |
|---|---|---|---|
| 0 | 2 | 5 | 12 |
| 3 | 8 | 4 | 9 |
| 2 | 7 | 3 | 9 |
| 1 | 3 | 0 | 3 |
| 5 | 13 | 4 | 10 |

(a) Find the 5-number summary and the mean.

(b) Draw a box plot.

(c) Draw a scatter plot of the data.

(d) Use regression to find the equation for the line of best fit.

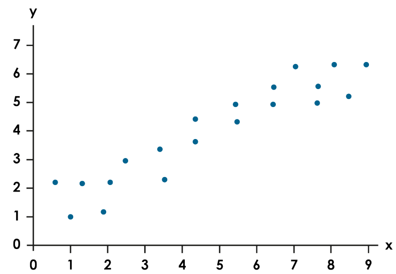
(e) Draw the line of best fit on the scatter plot.

(f) What is the slope of the line of best fit? What does it represent?

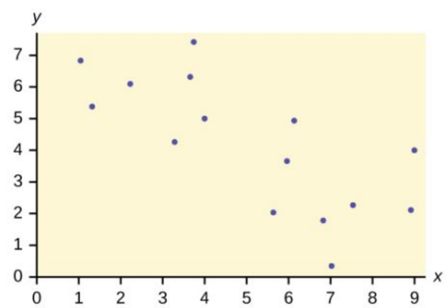
(g) What is the y-intercept of the line of best fit? What does it represent?

**Problem 5.25** Classify the following scatter plots: Are they linear? Strongly or weakly correlated? Positive or negative correlation? Explain in words.

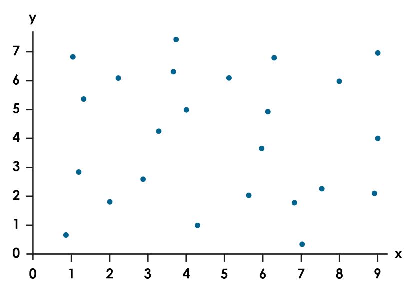
(a)




(b)




(c)



**Problem 5.26** Amelia plays basketball for her high school. She wants to improve to play at the college level and notices that the number of points she scores in a game goes up in response to the number of hours she practices her jump shot each week. She records the following data.

| X (hours practicing jump shot) | Y (points scored in a game) |
|---|---|
| 5 | 15 |
| 7 | 22 |
| 9 | 28 |
| 10 | 31 |
| 11 | 33 |
| 12 | 36 |

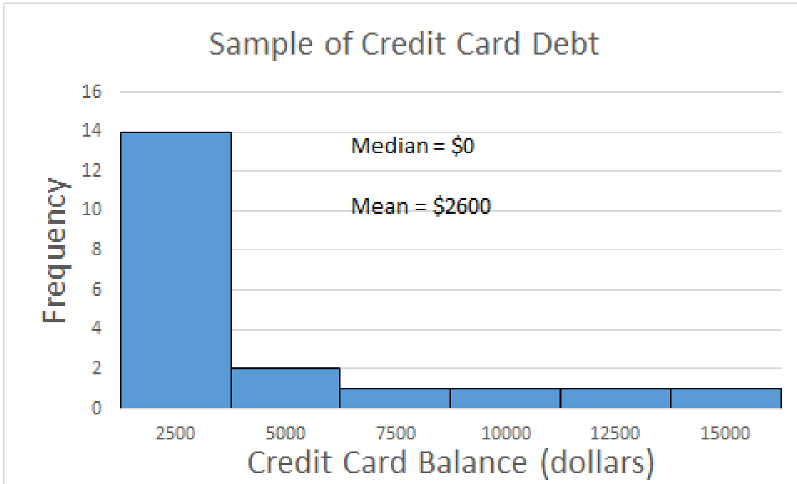
Construct the scatter plot and state if what Amelia thinks appears to be true.

**Problem 5.27** SCUBA divers have maximum dive times they cannot exceed when going to different depths. Draw a scatter plot with the least-squares regression line and predict the maximum dive time for 110 feet.

| X (Depth in Feet) | Y (Maximum Dive Time) |
|---|---|
| 50 | 80 |
| 60 | 55 |
| 70 | 45 |
| 80 | 35 |
| 90 | 25 |
| 100 | 22 |

The data in the table show different depths with the maximum dive times in minutes.

**Problem 5.28** Imagine that a sample of individuals report the balance on their credit card. The histogram summarizes this report:




(a) How many individuals are included in the sample?

(b) Which bin contains the mean credit card balance?

(c) The median for this sample is \$0. Explain what it means for the median to equal zero dollars.

(d) What is the mode for this sample?

(e) The mean of the sample is \$2600. Create a data set with a median of \$0 and a mean of \$2600 that matches the histogram.

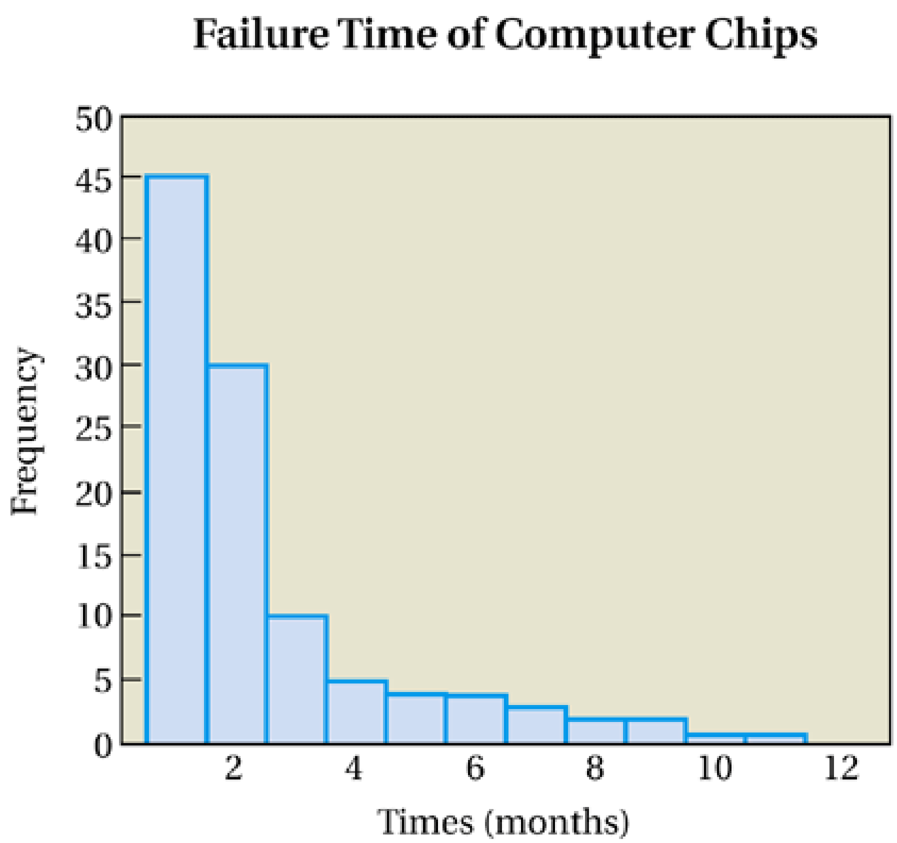
**Problem 5.29** A teacher makes the following list of the grades she gave to her 25 students on an essay:

$$A,C,C,B,C,D,C,C,F,D,C,C,C,B,B,A,B,D,B,A,A,B,F,C,B.$$

Complete the following table

| Grade | Frequency | Relative Frequency | Cumulative Frequency |
|---|---|---|---|
| A | 4 | 16% | |
| | | | |
| | | | |
| | | | |
| | | | |
| Total | | | |

**Problem 5.30** The time until failure for a sample of 107 computer chips that failed are shown below:




(a) Find the median, mode, and mean failure time of the computer chips in the sample.

(b) Comment on the shape, center, and spread of the graph.

(c) Write an interpretation of why the graph looks like it does.

**Problem 5.31** Give an example of:

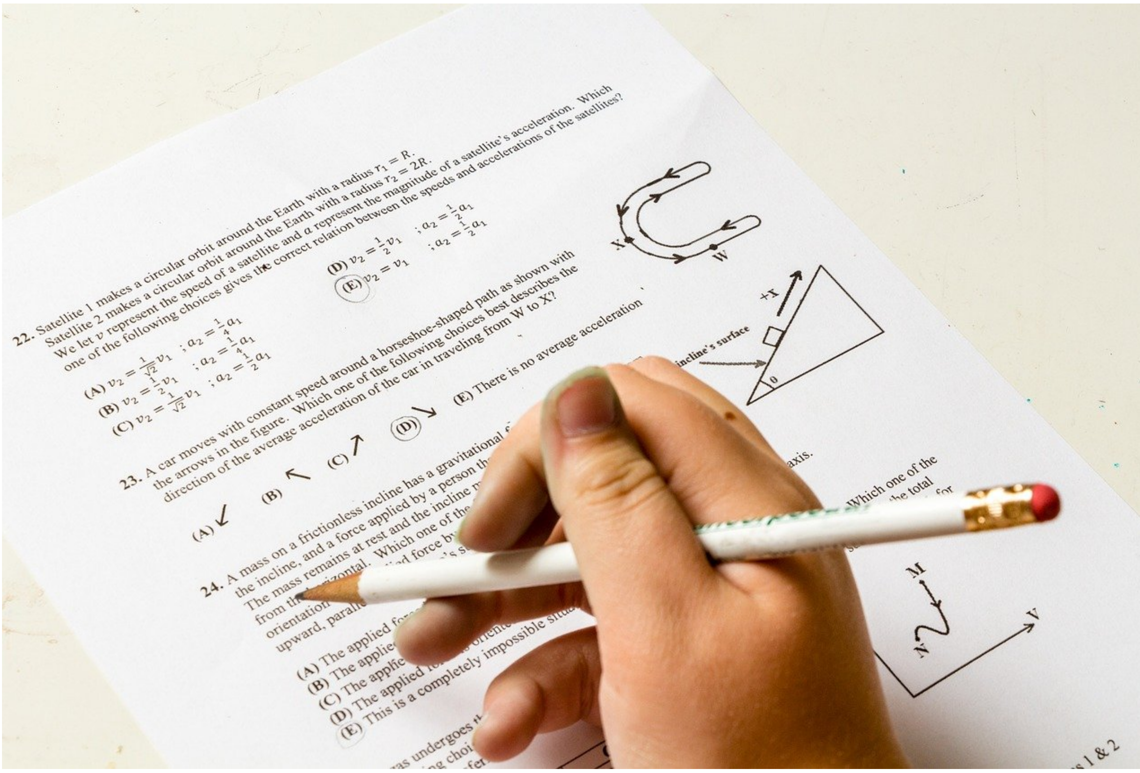
(a) Two variables correlated but without causation between them.

(b) Two causal situations that are not obviously correlated.

(c) A clear causal and correlation relation.

## 5.4 Projects

### 5.4.1 Grade Rubric and Test 2




The grades for Test #2 in a class are:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 84 | 73 | 46 | 56 | 69 | 23 | 68 | 43 | 64 | 74 |
| 31 | 70 | 64 | 41 | 42 | 56 | 15 | 57 | 81 | 39 |
| 58 | 56 | 59 | 68 | 59 | 63 | 53 | 75 | 59 | 43 |
| 91 | 76 | 53 | 64 | 77 | 34 | 73 | 61 | 56 | 62 |

(a) What type of questions can you answer with this data?

(b) Would an analysis of this data give you and experimental or observational study?

(c) Calculate the following:

1. Mean

2. Median

3. Mode

4. Range

5. Standard deviation

(d) The normal distribution z-scores for a data values are given by:

$$z = \frac{\text{data value} - \text{mean}}{\text{SD}}$$

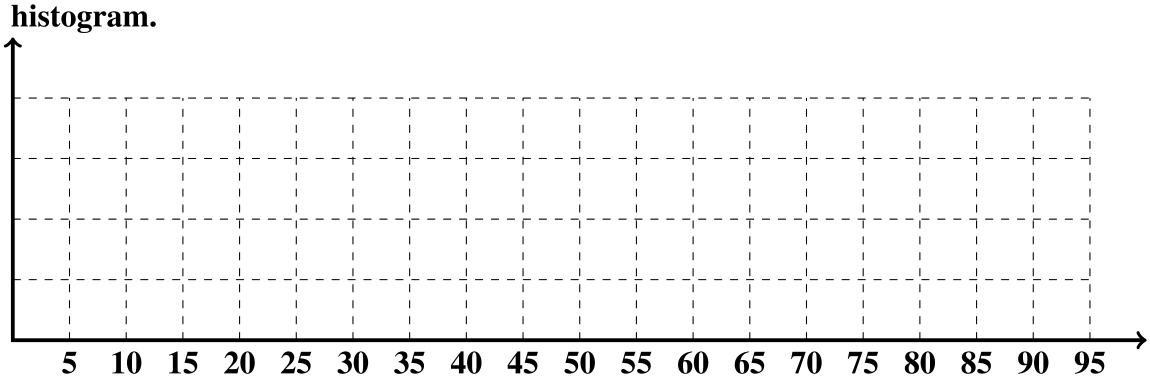
What $x$-value would have a z-score of $+1$?, $-1$?, $+2$?, $-2$? 0?

| | | | | | |
|---|---|---|---|---|---|
| $x$-value | | | | | |
| $z$-score | $-2$ | $-1$ | 0 | $+1$ | $+2$ |

(e) Create a data frequency table. First, choose intervals or bins into which the data values will fall. Typically, grades use bins with intervals of the length of 10.

| **Bins** | $0-9$ | $10-19$ | $20-29$ | $30-39$ | $40-49$ | $50-59$ | $60-69$ | $70-79$ | $80-89$ | $90+$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **Freq** | | | | | | | | | | |

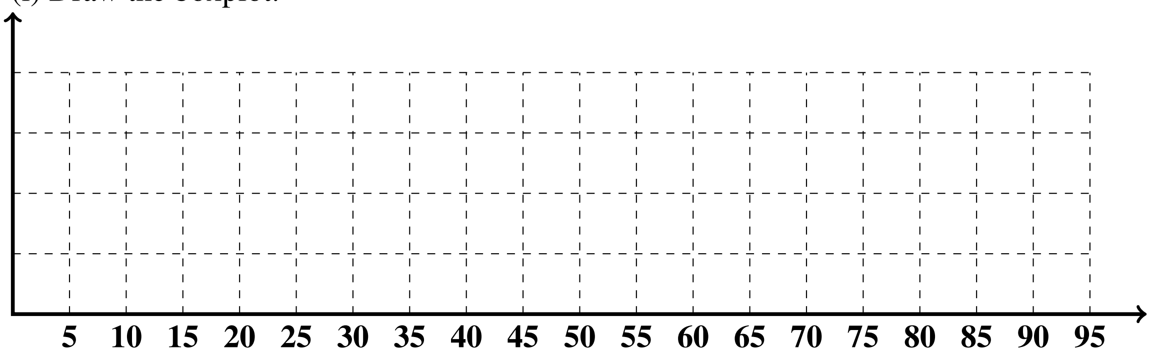
(f). Draw a histogram with bins on the $x-$axis. Put a tick mark on the $x$-axis for each of the $x$-values found in question #2. **Be sure to label each axis and title the histogram.**



(g) Group the data into letter grades, A, B, C, D, and F. Use the letter grades to create a pie chart.

(h) Give the five-number summary for the numerical scores on the test.

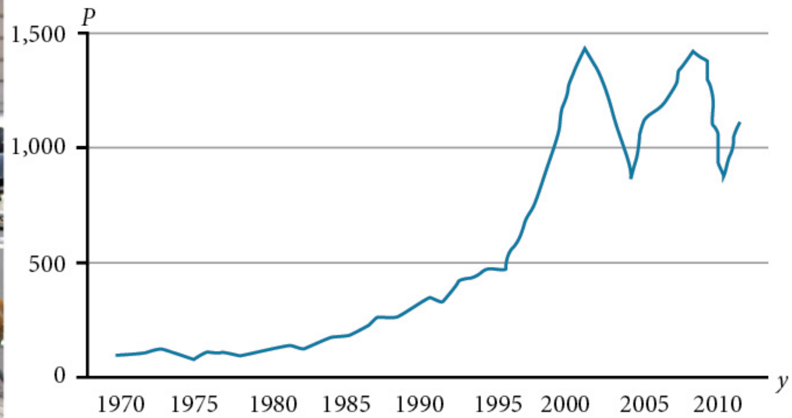
(i) Draw the boxplot.




(j) Suppose the letter grades are assigned following the below table based on $z$-scores

| | $z$-value | $x$-value . |
|---|---|---|
| **A** | $z \geq 2$ | |
| **B** | $1.5 \leq z < 2$ | |
| **C** | $1 \leq z < 1.5$ | |
| **D** | $0 \leq z < 1$ | |

- How many students did get an A? How many did get C's? How many did fail?
- Do you like or dislike this system? Explain.
- What systems would you use? Explain.




# 6. Basic Mathematical Modeling

Toward the end of the twentieth century, the values of stocks of Internet and technology companies rose dramatically. As a result, the Standard and Poor's stock market average also rose. The graph above tracks the value of that initial investment of just under \$100 over the 40 years. It shows that an investment worth less than \$500 until about 1995 skyrocketed up to about \$1100 by the beginning of 2000. That five-year period became known as the "dot-com bubble" because so many Internet startups were formed. As bubbles tend to do, though, the dot-com bubble eventually burst. Many companies grew too fast and then suddenly went out of business. The result caused the sharp decline shown on the graph beginning at the end of 2000.

In this example, there is a definite relationship between the year and the stock market average. For any year we choose, we can determine the corresponding value of the stock market average. This chapter will explore these kinds of relationships and their properties.

In this chapter, we aim to introduce the basic notions of mathematical modeling. First, we overview the concept of functions and some of their representations; after that, we dive into some of the most common and naturally occurring ways of modeling phenomena with linear and exponential modeling, finishing with quadratics and some of the elements that they can represent.

Parts of this chapter for the topics of functions, linear and exponential modeling have been extracted from the textbook **Open Stax- College Algebra** (see [Abr15]). To see the textbook, check or download it for free at `https://openstax.org/details/books/college-algebra`.

## 6.1 Introduction to Functions

A jetliner changes altitude as its distance from the starting point of a flight increases. The weight of a growing child increases with time. In each case, one quantity depends on another. There is a relationship between the two quantities that we can describe, analyze, and use to make predictions. In this section, we will explore such relationships.

### 6.1.1 Determining Whether a Relation Represents a Function

A relation is a set of ordered pairs. The set of the first components of each ordered pair is called the **domain** and the set of the second components of each ordered pair is called the **range**.

■ **Example 6.1 — Pairs.** Consider the following set of ordered pairs. The first numbers in each pair are the first five natural numbers. The second number in each pair is twice that of the first.

$$\{(1,2),(2,4),(3,6),(4,8),(5,10)\}$$

The domain is $\{1,2,3,4,5\}$. The range is $\{2,4,6,8,10\}$. ■

Note that each value in the domain is known as an input value or independent variable and is often labeled with the lowercase letter $x$. Each value in the range is also known as an output value or dependent variable and is often labeled lowercase letter $y$.

A function f is a relation that assigns a single value in the range to each value in the domain. In other words, no x-values are repeated. For our example that relates the first five natural numbers to numbers double their values, this relation is a function because each element in the domain, $\{1,2,3,4,5\}$, is paired with exactly one element in the range, $\{2,4,6,8,10\}$.

■ **Example 6.2 — Even-Odd Pairs.** Now let's consider the set of ordered pairs that relate the terms "even" and "odd" to the first five natural numbers. It would appear as

$$\{(\text{odd},1),(\text{even},2),(\text{odd},3),(\text{even},4),(\text{odd},5)\}$$

Notice that each element in the domain, even,odd is not paired with exactly one element in the range, $\{1,2,3,4,5\}$. For example, the term "odd" corresponds to three values from the range, $\{1,3,5\}$ and the term "even" corresponds to two values from the range, $\{2,4\}$.

This violates the definition of a function, so this relation is not a function. ■

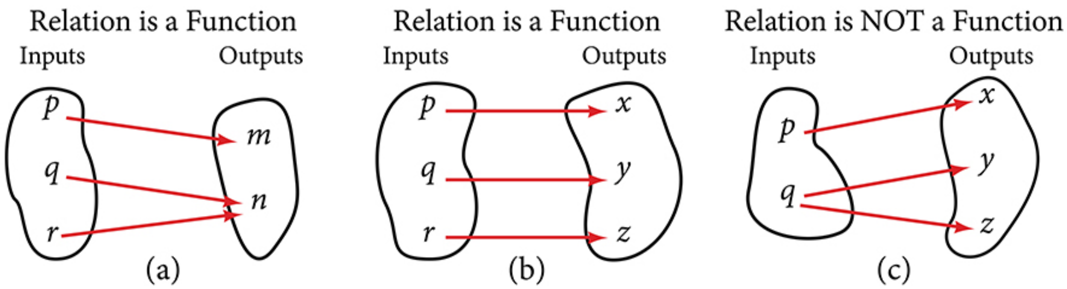



Figure 6.1: (a) This relationship is a function because each input is associated with a single output. Note that input q and r both give output n. (b) This relationship is also a function. In this case, each input is associated with a single output. (c) This relationship is not a function because input q is associated with two different outputs.

**Definition 6.1.1** A function is a relation in which each possible input value leads to exactly one output value. We say, "the output is a function of the input."

The input values make up the domain, and the output values make up the range.

**Definition 6.1.2** Given a relationship between two quantities, determine whether the relationship is a function.

- Identify the input values.
- Identify the output values.
- If each input value leads to only one output value, classify the relationship as a function. If any input value leads to two or more outputs, do not classify the relationship as a function.

■ **Example 6.3 — Percent grade.** In a particular math class, the overall percent grade corresponds to a grade point average. Is grade point average a function of the percent grade? Is the percent grade a function of the grade point average?

| Percent Grade | 0-56 | 57-61 | 62-66 | 67-71 | 72-77 | 78-86 | 87-91 | 92-100 |
|---|---|---|---|---|---|---|---|---|
| Grade point average | 0.0 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 |

Table 6.1: Shows a possible rule for assigning grade points.

**Solution:**

There is an associated grade point average for any grade percentage earned, so the grade point average is a function of the percent grade. In other words, if we input the percent grade, the output is a specific grade point average.

The grading system has a range of percent grades that correspond to the same grade point average. For example, students who receive a grade point average of 3.0 could have a variety of percent grades ranging from 78 to 86. Thus, percent grade is not a function of grade point average. ∎

∎ **Example 6.4 — The coffee shop.** The coffee shop menu, shown below, consists of items and their prices.

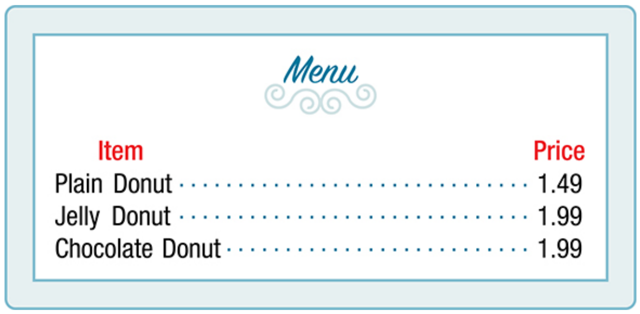



Figure 6.2: The Menu of Coffee Shop

1. Is price a function of the item?
2. Is the item a function of the price?

**Solution:**

1. Let's begin by considering the input as the items on the menu. The output values are then the prices. Each item on the menu has only one price, so the price is a function of the item.
2. Two items on the menu have the same price. If we consider the prices to be the input values and the items to be the output, the same input value could have more than one associated output. See the image below.

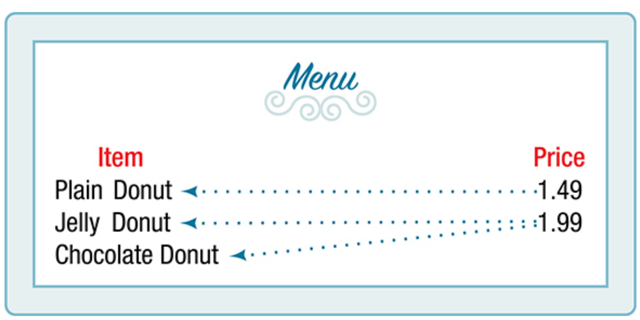



Figure 6.3: The menu relationship from price to the item.

Therefore, the item is not a function of price.

∎

### 6.1.2 Using Function Notation

Once we determine that a relationship is a function, we must display and define the functional relationships to understand and use them. Sometimes we can program them into computers. There are various ways of representing functions. A standard function notation is one representation that facilitates working with functions.

To represent "height is a function of age," we start by identifying the descriptive variables $h$ for height and $a$ for age. The letters $f, g$, and $h$ are often used to represent functions just as we use $x, y$, and $z$ to represent numbers and $A, B$, and $C$ to represent sets.

- $h$ is $f$ of $a$. (We name the function $f$; height is a function of age.)
- $h = f(a)$. (We use parentheses to indicate the function input.)
- $f(a)$ (We name the function $f$;the expression is read as "$f$ of $a$.")

Remember, we can use any letter to name the function; the notation $h(a)$ shows us that $h$ depends on $a$. The value $a$ must be put into the function $h$ to get a result. The parentheses indicate that age is input into the function; they do not indicate multiplication.

We can also give an algebraic expression as the input to a function. For example $f(a+b)$ means "first add $a$ and $b$, and the result is the input for the function $f$." We must perform operations in this order to obtain the correct result.

**Definition 6.1.3** The notation $y = f(x)$ defines a function named $f$. This is read as "$y$ is a function of $x$." The letter $x$ represents the input value, or independent variable. The letter $y$, or $f(x)$, represents the output value, or dependent variable.

■ **Example 6.5 — Using function notation for days in a month.** Use function notation to represent a function whose input is the name of a month and output is the number of days in that month. Assume that the domain does not include leap years. ■

**Solution:** The number of days in a month is a function of the name of the month, so if we name the function $f$, we write **days=f(month)** or $d = f(m)$. The name of the month is the input to a "rule" associating a specific number (the output) with each input.

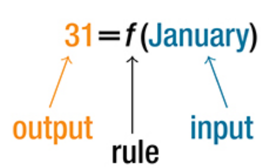



Figure 6.4: The days in function of the month.

For example, **f(March)=31**, because March has 31 days. The notation $d = f(m)$ reminds us that the number of days, d (the output), is dependent on the month, m (the input).

■ **Example 6.6 — Interpreting Function Notation.** A function $N = f(y)$ gives the number of police officers, $N$, in a town in year $y$. What does $f(2005) = 300$ represent?

**Solution:** When we read $f(2005) = 300$, we see that the input year is 2005. The value for the output, the number of police officers (N), is 300. Remember, $N = f(y)$. The statement $f(2005) = 300$ tells us that in 2005, there were 300 police officers in the town. ■

### 6.1.3 Representing Functions Using Tables

A standard method of representing functions is in the form of a table. The table rows or columns display the corresponding input and output values. In some cases, these values represent all we know about the relationship; other times, the table provides a few select examples from a complete relationship.

■ **Example 6.7 — Month and Days.** Consider the table:

| **Month number, m (input)** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Days in month, D (output)** | 31 | 28 | 31 | 30 | 31 | 30 | 31 | 31 | 30 | 31 | 30 | 31 |

Table 6.2: Days by month in table.

Table 6.2 lists the input number of each month (January = 1, February = 2, and so on) and the output value of the number of days in that month. This information represents all we know about the months and days of a given year (that is not a leap year). In this table, we define a days-in-a-month function $f$ where $D = f(m)$ identifies months by an integer rather than by name. ■

■ **Example 6.8 — Input-Output.** Now

| **n** | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Q** | 8 | 6 | 7 | 6 | 8 |

Table 6.3: Input $n$ and output $Q$

Table 6.3 defines a function $Q = g(n)$. Remember, this notation tells us that g is the name of the function that takes the input **n** and gives the output **Q**. ■

■ **Example 6.9 — Age and Height.** Finally

| **Age in years, a (input)** | 5 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|
| **Height in inches, h (output)** | 40 | 42 | 44 | 47 | 50 | 52 | 54 |

Table 6.4: Age in terms of height

Table 6.4 displays the age of children in years and their corresponding heights. Right away, we can see that this table does not represent a function because the same input value, five years, has two different output values, 40 in. and 42 in. ■

■ **Example 6.10 — What table is the function?.** Consider the following tables:

1.

| Input | Output |
|---|---|
| 2 | 1 |
| 5 | 3 |
| 8 | 6 |

2.

| Input | Output |
|---|---|
| -3 | 5 |
| 0 | 1 |
| 4 | 5 |

3.

| Input | Output |
|---|---|
| 1 | 0 |
| 5 | 2 |
| 5 | 4 |

which tables represents a function (if any)? ■

**Solution:**

Tables 1 and 2 define functions. In both, each input value corresponds to exactly one output value. Table 3 does not specify a function because the input value of 5 corresponds to two different output values. When a table represents a function, corresponding input and output values can be specified using function notation.

The function represented in Table 1 can be defined by writing.

$$f(2) = 1, f(5) = 3, f(8) = 6$$

Similarly, the statements.

$$g(-3) = 5, g(0) = 1, g(4) = 5$$

represent the function in Table 2.

We cannot express Table 3 values similarly because it does not represent a function.

### 6.1.4 Finding Input and Output Values of a Function

When we know an input value and want to determine the corresponding output value for a function, we evaluate the function. Evaluating will always produce one result because each input value of a function corresponds to exactly one output value.

When we know an output value and want to determine the input values that would produce that output value, we set the output equal to the function's formula and solve for the input. Solving can create more than one solution because different input values can have the same output value.

#### Evaluation of Functions in Algebraic Forms

When we have a function in formula form, it is usually a simple matter to evaluate the function. For example, the function $f(x) = 5 - 3x^2$ can be evaluated by squaring the input value, multiplying by 3, and then subtracting the product from 5.

**Definition 6.1.4** Given the formula for a function, evaluate;

- Substitute the input variable in the formula with the value provided.
- Calculate the result.

■ **Example 6.11 — Evaluating Functions at Specific Values.** Evaluate $f(x) = x^2 + 3x - 4$ at:

- 2
- $-3$
- $a$

**Solution:**

Replace the $x$ in the function with each specified value.

- Because the input value is a number, 2, we can use simple algebra to simplify.

$$f(2) = 2*2 + 3*(2) - 4 = 4 + 6 - 4 = 6$$

- Because the input value is a number, $-3$, we can use simple algebra to simplify.

$$f(2) = (-3)*(-3) + 3*(-3) - 4 = 9 - 9 - 4 = -4$$

- In this case, the input value is a letter so we cannot simplify the answer any further.

$$f(a) = a^2 + 3a - 4$$

■

■ **Example 6.12 — Evaluating functions.** Given the function

$$h(p) = p^2 + 2p$$

evaluate $h(4)$

**Solution:**

To evaluate $h(4)$, we substitute the value 4 for the input variable $p$ in the given function.

$$h(p) = p^2 + 2p$$

$$h(4) = 4^2 + 2*4 = 16 + 8 = 24$$

■

### Evaluating a Function Given in Tabular Form

As we saw above, we can represent functions in tables. Conversely, we can use information in tables to write functions and evaluate functions using the tables. For example, how well do our pets recall the fond memories we share with them? There is an urban legend that a goldfish has a memory of 3 seconds, but this is just a myth. Goldfish can remember up to 3 months, while the beta fish has a memory of up to 5 months. And while a puppy's memory span is no longer than 30 seconds, the adult dog can remember for 5 minutes; this is meager compared to a cat, whose memory span lasts 16 hours.

The function that relates the type of pet to the duration of its memory span is more easily visualized with a table. See

| Pet | Memory span in hours |
|---|---|
| Puppy | 0.008 |
| Adult dog | 0.083 |
| Cat | 16 |
| Goldfish | 2160 |
| Beta fish | 3600 |

Figure 6.5: Memory span in hours in function of a pet

Evaluating a function in table form may be more valuable than using equations. Here let us call the function P. The function's domain is the pet type, and the range is a real number representing the number of hours the pet's memory span lasts. We can evaluate the function P at the input value of "goldfish." We would write P(goldfish)=2160. Notice that, to evaluate the function in table form, we identify the input value and the corresponding output value from the pertinent row of the table. The tabular form for function P seems ideally suited to this function, more so than writing it in paragraph or function form.

**Definition 6.1.5** Given a function represented by a table, identify specific output and input values.

- Find the given input in the row (or column) of input values.
- Identify the corresponding output value paired with that input value.
- Find the given output values in the row (or column) of output values, noting every time that output value appears.
- Identify the input value(s) corresponding to the given output value.

■ **Example 6.13 — Evaluating and Solving a Tabular Function.** Using the following table.

| $n$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $g(n)$ | 8 | 6 | 7 | 6 | 8 |

Calculate:

- $g(3)$.
- Solve $g(n) = 6$.

■

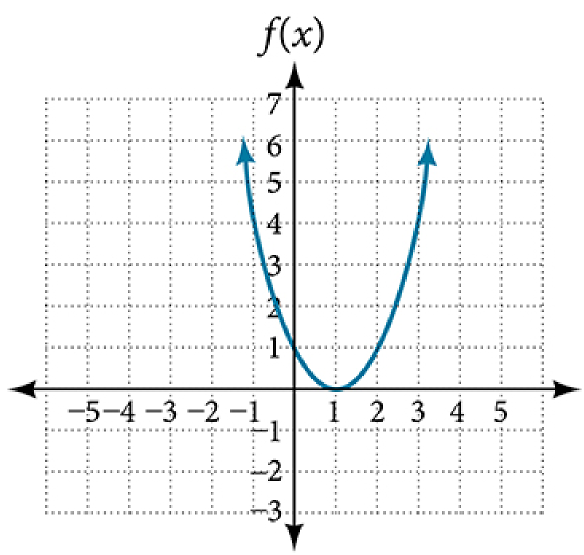
**Solution:**

- Evaluating $g(3)$ means determining the output value of the function $g$ for the input value of $n = 3$. The table output value corresponding to $n = 3$ is 7, so $g(3) = 7$.
- Solving $g(n) = 6$ means identifying the input values, $n$, that produce an output value of 6. The table shows two solutions: 2 and 4.

### Finding Function Values from a Graph

Evaluating a function using a graph also requires finding the corresponding output value for a given input value; only, in this case, we see the output value by looking at the graph. Solving a function equation using a diagram requires finding all instances of the given output value on the chart and observing the corresponding input value(s).

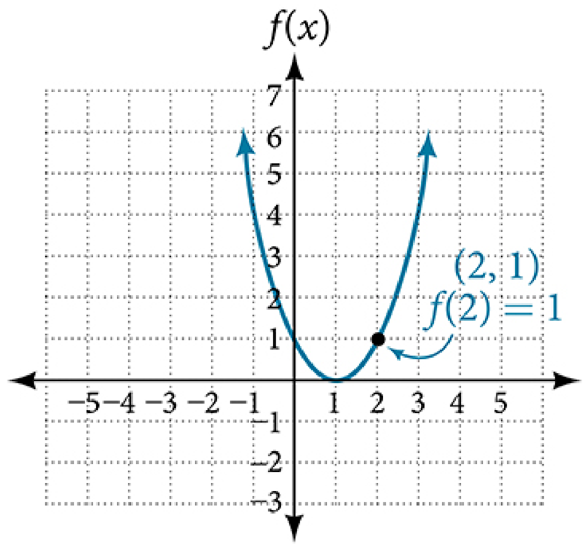
■ **Example 6.14 — Reading Function Values from a Graph.** Given the graph Evaluate



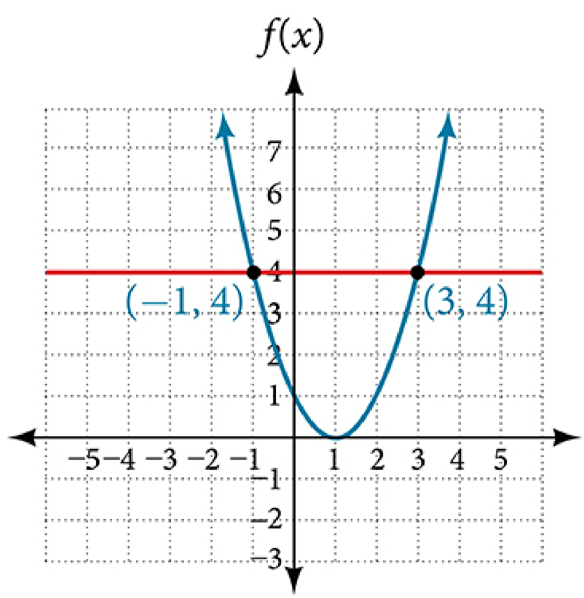

$f(2)$ and find where $f(x) = 4$. ■

**Solution:**

To evaluate $f(2)$, locate the point on the curve where $x = 2$, then read the $y$-coordinate of that point. The point has coordinates $(2,1)$, so $f(2) = 1$. See




To solve $f(x) = 4$, we find the output value 4 on the vertical axis. Moving horizontally along the line $y = 4$, we locate two points of the curve with output value 4: $(-1,4)$ and $(3,4)$. These points represent the two solutions to $f(x) = 4$: $-1$ or 3. This means $f(-1) = 4$ and $f(3) = 4$, or when the input is $-1$ or 3, the output is 4. See



### 6.1.5 More in Domain and Range

In section 6.1, we were introduced to the concepts of domain and range. In this section, we will practice determining domains and ranges for specific functions. Keep in mind that, in determining domains and ranges, we need to consider what is physically possible or meaningful in real-world examples, such as ticket sales and year (see the graph below).

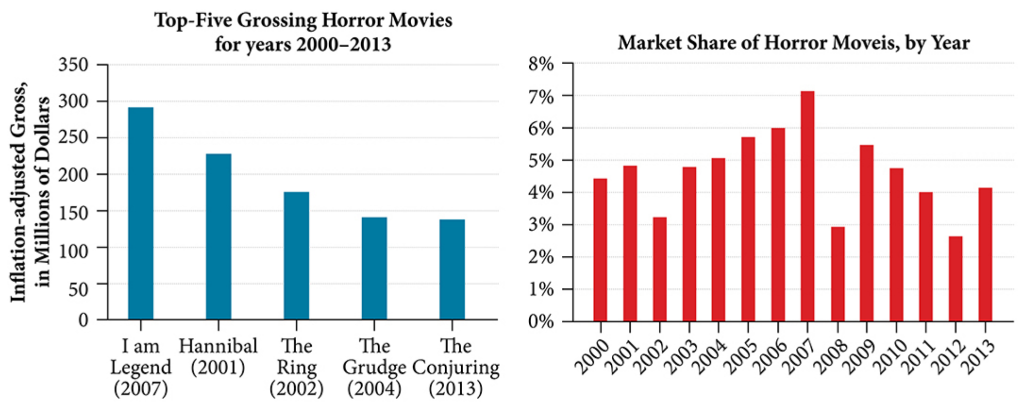



Figure 6.6: Horror movies numbers

Figure 6.6 how the amount, in dollars, each of those movies grossed when they were released and the ticket sales for horror movies in general by year. Notice that we can use the data to create a function of each film's Box Office or the total ticket sales for all horror movies by year.

We also need to consider what is mathematically permitted. For example, we cannot include any input value that leads us to take an even root of a negative number if the domain and range consist of real numbers. Or, in a function expressed as a formula, we cannot include any input value in the domain that would lead us to divide by 0.

We can visualize the domain as a "holding area" that contains "raw materials" for a "function machine" and the range as another "holding area" for the machine's products. See

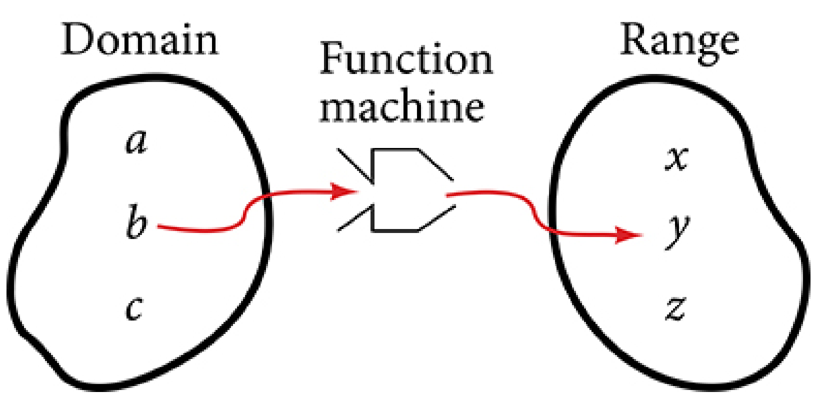



Figure 6.7: The function machine

We can write the domain and range in interval notation, which uses values within brackets to describe a set of numbers. In interval notation, we use a square bracket [ when the set includes the endpoint and a parenthesis ( to indicate that the endpoint is not included or the interval is unbounded. For example, if a person has \$100 to spend, they would need to express the interval that is more than 0 and less than or equal to 100 and write $(0, 100]$.

Before we begin, let us review the conventions of interval notation:

- The smallest number from the interval is written first.
- The largest number in the interval is written second, following a comma.
- Parentheses, ( or ), are used to signify that an endpoint value is not included, called exclusive.
- Brackets, [ or ], is used to indicate that an endpoint value is included, called inclusive.

| Inequality | Interval Notation | Graph on Number Line | Description |
|---|---|---|---|
| $x > a$ | $(a, \infty)$ |   | *x* is greater than *a* |
| $x < a$ | $(-\infty, a)$ |   | *x* is less than *a* |
| $x \geq a$ | $[a, \infty)$ |   | *x* is greater than or equal to *a* |
| $x \leq a$ | $(-\infty, a]$ |   | *x* is less than or equal to *a* |
| $a < x < b$ | $(a, b)$ |   | *x* is strictly between *a* and *b* |
| $a \leq x < b$ | $[a, b)$ |   | *x* is between *a* and *b*, to include *a* |
| $a < x \leq b$ | $(a, b]$ |   | *x* is between *a* and *b*, to include *b* |
| $a \leq x \leq b$ | $[a, b]$ |   | *x* is between *a* and *b*, to include *a* and *b* |

■ **Example 6.15 — Ordered Pairs.** Find the domain of the following function:

$$\{(2,10),(3,10),(4,20),(5,30),(6,40)\}$$

.

**Solution:** First, identify the input values. The input value is the first coordinate in an ordered pair. There are no restrictions, as the ordered pairs are simply listed. The domain is the set of the first coordinates of the ordered pairs.

$$\{2,3,4,5,6\}$$

■

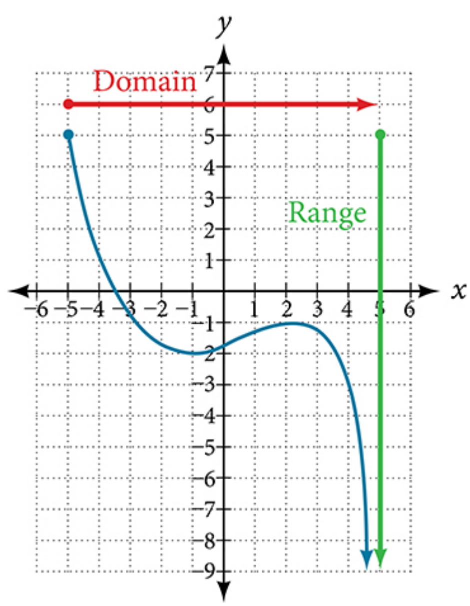
■ **Example 6.16 — Input-Output.** Find the Domain and Range of the function given by the table.

| **n** | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Q** | 8 | 6 | 7 | 6 | 8 |

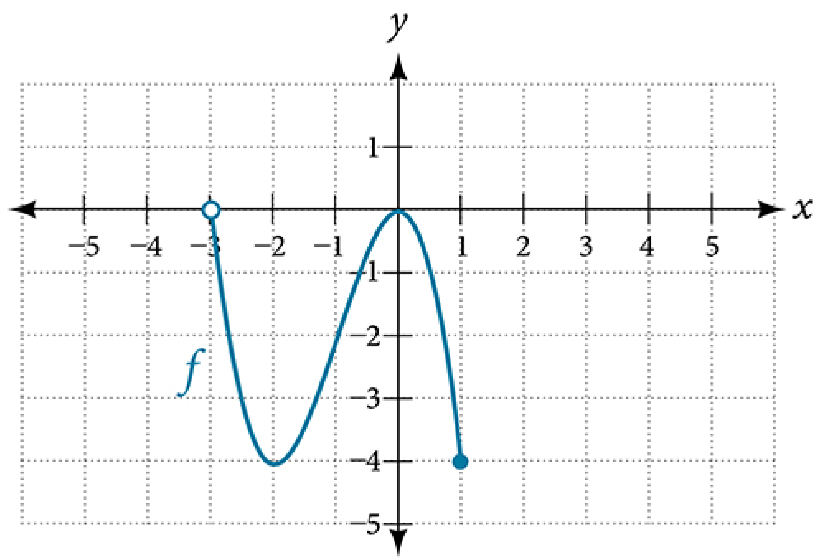
**Solution:** First identify the input values $n$ and the output values $Q$. Then the domain is $\{1,2,3,4,5\}$ and the range $\{6,7,8\}$. ■

### Finding Domain and Range from Graphs

Another way to identify the domain and range of functions is using graphs. Because the domain refers to the set of possible input values, the domain of a graph consists of all the input values shown on the x-axis. The range is the set of possible output values shown on the y-axis. Keep in mind that if the graph continues beyond the portion of the chart, we can see that the domain and range may be greater than the visible values. See

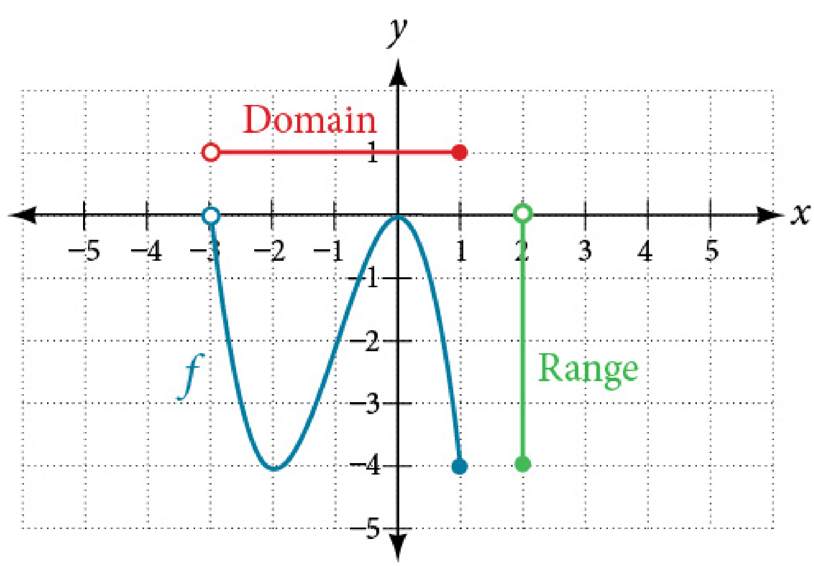


■ **Example 6.17 — Domain and Range from graph.** Find the domain and range of the function f whose graph is shown in




**Solution:**

Observe that the horizontal extent of the graph is $-3$ to 1, so the domain of f is $(-3,1]$. Similarly, the vertical extent of the graph is 0 to $-4$, so the range is $[-4,0)$. See notice that




the function has no interruptions therefore making our answer two single intervals. Graphs with multiple pieces may require union of multiple intervals. ■

### 6.1.6 Exercises 6.1

**Problem 6.1** What is the difference between a relation and a function?

**Problem 6.2** What is the difference between the input and the output of a function?

**Problem 6.3** For the following exercises, determine whether the relation represents a function.

(i) $\{(a,b),(c,d),(a,c)\}$

(ii) $\{(a,b),(b,c),(c,c)\}$

(iii) $\{(a,a),(a,c),(c,c)\}$

**Problem 6.4** For the following exercises, evaluate the function $f$ at the indicated values

$$f(-3), f(2), f(1), f(a), f(0).$$

(a) $f(x) = 2x - 5$

(b) $f(x) = -5x^2 + 2x - 1$

(c) $f(x) = \sqrt{2-x} + 5$

(d) $f(x) = \frac{6x-1}{5x+2}$

(e) $f(x) = 5$

**Problem 6.5** Given the function $f(x) = 8 - 3x$

(a) Evaluate $f(2)$. (b) Solve $f(x) = -1$

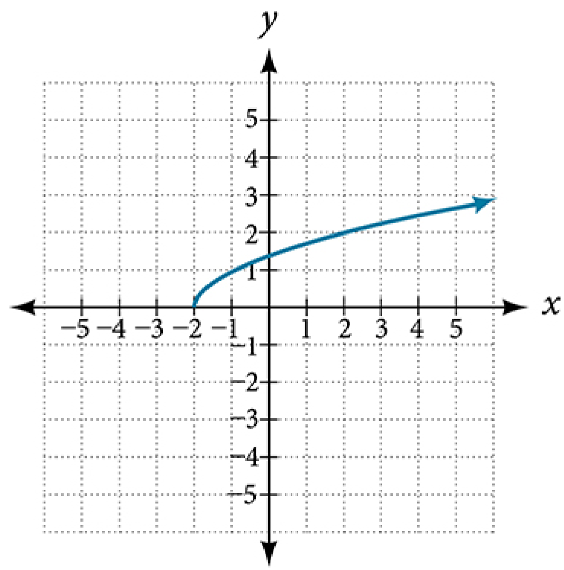
**Problem 6.6** Given the function $p(c) = c^2 + c$

(a) Evaluate $p(-3)$. (b) Solve $p(c) = 2$

**Problem 6.7** Given the function $f(x) = \sqrt{x+2}$

(a) Evaluate $f(7)$. (b) Solve $f(x) = 4$

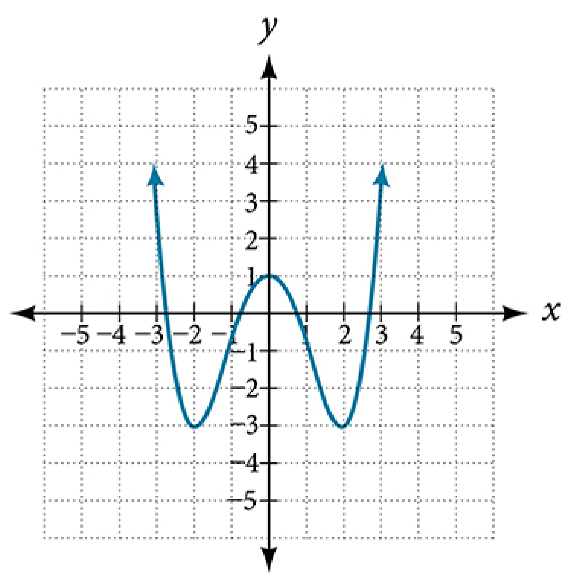
**Problem 6.8** Given the following graph:




(a) Evaluate $f(1)$.

(b) Solve $f(x) = 3$

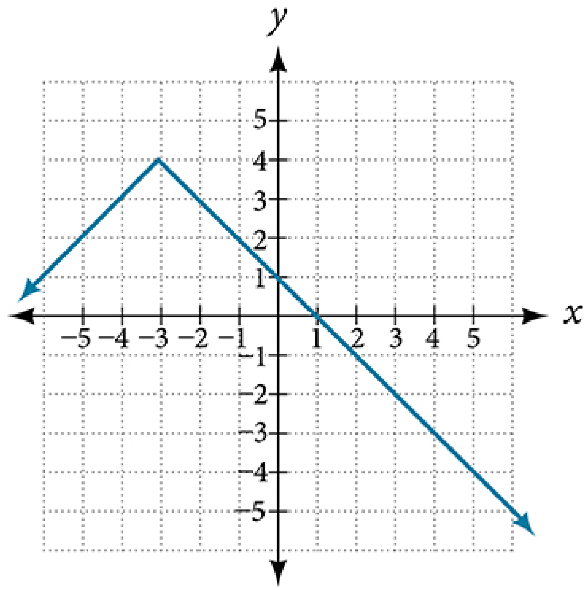
**Problem 6.9** Given the following graph:




(a) Evaluate $f(0)$.

(b) Solve $f(x) = -3$

**Problem 6.10** Given the following graph:




(a) Evaluate $f(4)$.

(b) Solve $f(x) = 1$

**Problem 6.11** For the following exercises, determine whether the relation represents a function.

- $\{(-1,1),(-2,2),(-3,3)\}$
- $\{(3,4),(4,5),(5,6)\}$
- $\{(2,5),(7,11),(15,8),(7,9)\}$

**Problem 6.12** The amount of garbage, $G$, produced by a city with a population $p$ is given by $G = f(p)$. $G$ is measured in tons per week, and $p$ is measured in thousands of people.

- The town of Tola has a population of 40,000 and produces 13 tons of garbage each week. Express this information in terms of the function f.
- Explain the statement's meaning for $f(5) = 2$.

**Problem 6.13** The following table shows the most medals won by any country in the given year of the Olympic Winter Games:

| **Year** | **Most medals won by a country** | **Country** |
|---|---|---|
| 1988 | 29 | Soviet Union |
| 1992 | 26 | Germany |
| 1994 | 26 | Norway |
| 1998 | 29 | Germany |
| 2002 | 36 | Germany |
| 2006 | 29 | Germany |
| 2010 | 37 | United States |
| 2014 | 29 | Russia |

1. Identify the independent and the dependent variables, and describe the domain and range.

2. Describe the function in words.

**Problem 6.14** Consider the following scenario: The amount of money in an individual's checking account sometimes increases (when she deposits her paychecks) and sometimes decreases (when she pays her bills).

1. Create a possible graph showing the amount of money in the checking account over two months. Explain why it looks as it does.

2. Based on your graph, is the amount of money in the checking account a function of the time of the year?

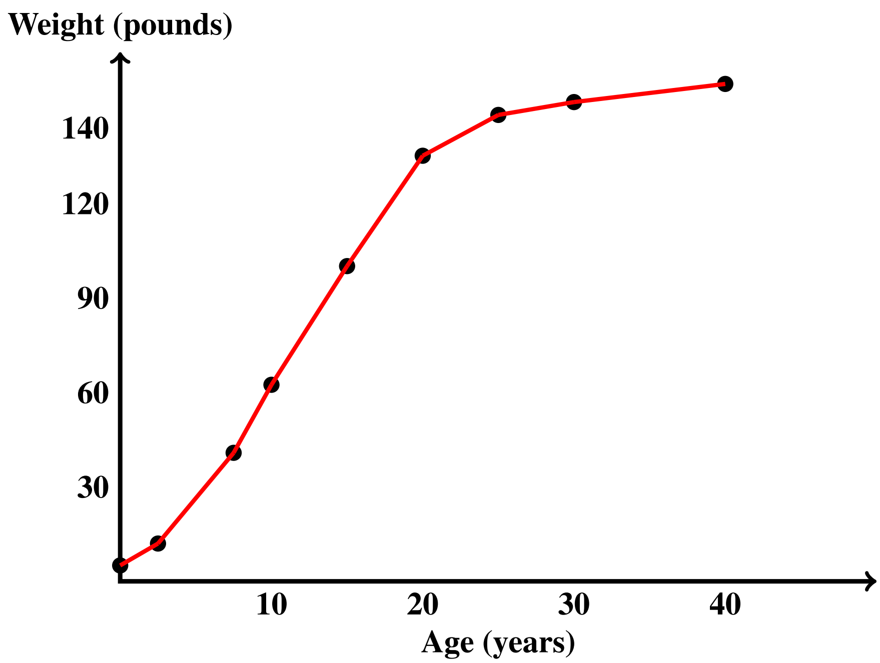
3. Based on your graph, is the time a function of the amount of money in the account?

**Problem 6.15** The number of cubic yards of dirt, $D$, needed to cover a garden with an area of square feet is given by $D = g(a)$.

(a) A garden with area 5000 $ft^2$ requires 50 $yd^3$ of dirt. Express this information in terms of the function $g$.

(b) Explain the meaning of the statement $g(100) = 1$.

**Problem 6.16** Consider the graph of the following function.




1. Identify the independent and the dependent variables, and describe the domain and range.

2. Describe the function in words.

**Problem 6.17** Given the following Table:

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $f(x)$ | 74 | 28 | 1 | 53 | 56 | 3 | 36 | 45 | 14 | 47 |

(a) Evaluate $f(3)$.

(b) Solve $f(x) = 1$.

(c) Evaluate $f(4)$.

(d) Solve $f(x) = 28$

### Problem 6.18

Determine if the relation in the table represents a function $y = f(x)$.

| | | | |
|---|---|---|---|
| $x$ | 5 | 10 | 15 |
| $y$ | 3 | 8 | 14 |

### Problem 6.19

Determine if the relation in the table represents a function $y = f(x)$.

| | | | |
|---|---|---|---|
| $x$ | 5 | 10 | 15 |
| $y$ | 3 | 8 | 8 |

**Problem 6.20** Why does the domain differ for different functions?

**Problem 6.21** The height $h$ of a projectile is a function of the time $t$ it is in the air. The height in feet for $t$ seconds is given by the function

$$h(t) = -16t^2 + 96t.$$

What is the domain of the function? What is the meaning of the domain in the context of the problem?

**Problem 6.22** The cost in dollars of making $x$ items is given by the function

$$C(x) = 10x + 500$$

Determine:

(a) The fixed cost is determined when zero items are produced. Find the fixed cost for this item.

(b) What is the cost of making 25 items?

(c) Suppose the maximum cost allowed is \$1500. What are the domain and range of the cost function, $C(x)$?

(d) Can we calculate the cost of 150 items with this function?

## 6.2 Linear Modeling

Figure 6.8: A bamboo forest in China (credit: "JFXie"/Flickr)

Imagine placing a plant in the ground one day and finding that it has doubled its height just a few days later. Although it may seem incredible, this can happen with certain bamboo species. These members of the grass family are the fastest-growing plants in the world. One species of bamboo has been observed to grow nearly 1.5 inches every hour. 6 In a twenty-four-hour period, this bamboo plant grows about 36 inches or an incredible 3 feet! A constant rate of change, such as the growth cycle of this bamboo plant, is a linear function.

Recall from Functions and Function Notation that a function is a relation that assigns to every element in the domain exactly one element in the range. Linear functions are a specific type of function we can use to model many real-world applications, such as plant growth over time. This chapter will explore linear functions, their graphs, and how to relate them to data.

Just as with the growth of a bamboo plant, many situations involve constant change over time. Consider, for example, the first commercial maglev train in the world, the Shanghai MagLev Train. It carries passengers comfortably for a 30-kilometer trip from the airport to the subway station in only eight minutes.

Suppose a maglev train travels a long distance and maintains a constant speed of 83 meters per second once it is 250 meters from the station. How can we analyze the train's distance from the station as a function of time? In this section, we will investigate a kind of function that is useful for this purpose and use it to explore real-world situations such as the train's distance from the station at a given time.


Figure 6.9: Shanghai MagLev Train (credit: "kanegen"/Flickr)

### 6.2.1 Representing Linear Functions

The function describing the train's motion is a linear function, defined as a function with a constant rate of change. There are several ways to represent a linear function, including word form, function notation, tabular form, and graphical form. We will describe the train's motion as a function using each method.

#### Representing a Linear Function in Word form

Let's begin by describing the linear function in words. For the train problem, we just considered, We may use the following word sentence to describe the function relationship.

- The train's distance from the station is a function of the time during which the train moved at a constant speed plus its original distance from the station when it began moving at a constant speed.

The speed is the rate of change. Recall that a rate of change measures how quickly the dependent variable changes concerning the independent variable. The rate of change for this example is constant, meaning it is the same for each input value. As the time (input) increases by 1 second, the corresponding distance (output) increases by 83 meters. The train began moving at this constant speed at 250 meters from the station.

### Representing a Linear Function in Function Notation

Another approach to representing linear functions is by using function notation. One example of function notation is an equation written in the slope-intercept form of a line, where $x$ is the input value, $m$ is the rate of change, and $b$ is the initial value of the dependent variable.

- **Equation form:** $y = mx + b$
- **Function form:** $f(x) = mx + b$

We can write a generalized equation to represent the motion of the train. In the example of the train, we might use the notation $D(t)$ where the total distance $D$ is a function of the time $t$. The rate, $m$, is 83 meters per second. The initial value of the dependent variable $b$ is the actual distance from the station, 250 meters.

$$D(t) = 83t + 250$$

### Representing a Linear Function in Tabular form

A third method of representing a linear function is through the use of a table. The relationship between the distance from the station and the time is represented in the figure 6.10 below:

1 second 1 second 1 second

| $t$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $D(t)$ | 250 | 333 | 416 | 499 |

83 meters 83 meters 83 meters

Figure 6.10: Tabular representation of the function $D$ showing selected input and output values

The table shows that the distance changes by 83 meters for every 1-second increase in time. Notice that the domain always sets a place and an essential role for things to make sense. Can the input in the previous example be any real number?

No. The input consists of non-negative real numbers. The input represents time, so while non-negative rational and irrational numbers are possible, negative real numbers are impossible for this example.

### Representing a Linear Function in Graphical Form

Another way to represent linear functions is visual, using a graph. We can use the function relationship from above, $D(t) = 83t + 250$, to draw a graph as represented in Figure 6.11. Notice the graph is a line. When we plot a linear function, the graph is always a line.

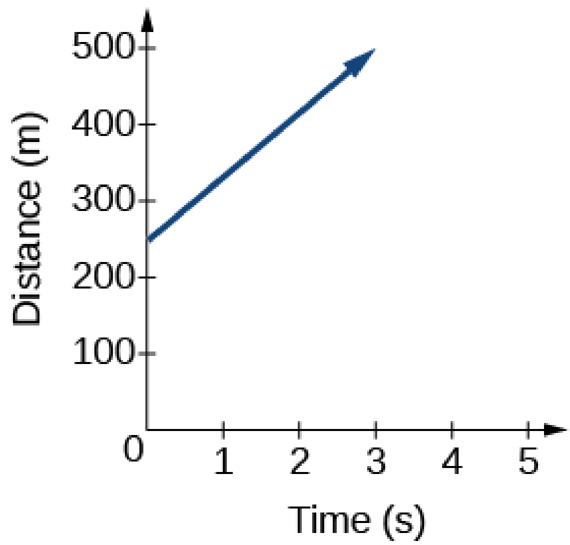



Figure 6.11: Tabular representation of the function $D$ showing selected input and output values

The rate of change, which is constant, determines the slant or slope of the line. The point at which the input value is zero is the line's vertical intercept or y-intercept. We can see from the graph that the y-intercept in the train example we just saw is (0,250) and represents the distance of the train from the station when it began moving at a constant speed.

Notice that the graph of the train example is restricted, but this is not always the case. Consider the graph of the line $f(x) = 2x + 1$. Ask yourself what numbers can be input into the function. In other words, what is the domain of the function? The domain is comprised of all real numbers because any number may be doubled and then have one added to the product.

**Definition 6.2.1** A linear function is a function whose graph is a line. We can write linear functions in the slope-intercept form of a line.

$$f(x) = mx + b$$

where $b$ is the initial or starting value of the function (when input, $x = 0$ ), and m is the constant rate of change, or slope of the function. The y-intercept is at $(0, b)$.

■ **Example 6.18 — Using a Linear Function to Find the Pressure on a Diver.** The pressure, $P$, in pounds per square inch (PSI) on the diver in Figure 6.12 depends upon her depth below the water surface, $d$, in feet.

Figure 6.12: (credit: Ilse Reijs and Jan-Noud Hutten)

This relationship may be modeled by the equation,

$$P(d) = 0.434d + 14.696.$$

Restate this function in words.

**Solution:**
To restate the function in words, we need to describe each part of the equation. The pressure as a function of depth equals four hundred thirty-four thousandths times depth plus fourteen and six hundred ninety-six thousandths.

The initial value, 14.696, is the pressure in PSI on the diver at a depth of 0 feet, which is the water's surface. The rate of change, or slope, is 0.434 PSI per foot; this tells us that the pressure on the diver increases by 0.434 PSI for each foot her depth increases.

■

### 6.2.2 Parts of a Linear Functions

#### Interpreting Slope as a Rate of Change

The problem provided the slope to us in the examples we have seen. However, we often need to calculate the slope given input and output values. Recall that given two values for the input, $x_1$ and $x_2$, and two corresponding values for the output, $y_1$ and $y_2$ which can be represented by a set of points, $(x_1, y_1)$ and $(x_2, y_2)$ we can calculate the slope $m$.

$$m = \frac{\text{change in output (rise)}}{\text{change in input (run)}} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \tag{6.1}$$

Note that in function notation we can obtain two corresponding values for the output $y_1$ and $y_2$ for the function $f$, $y_1 = f(x_1)$ and $y_2 = f(x_2)$, so we could equivalently write

$$m = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{6.2}$$

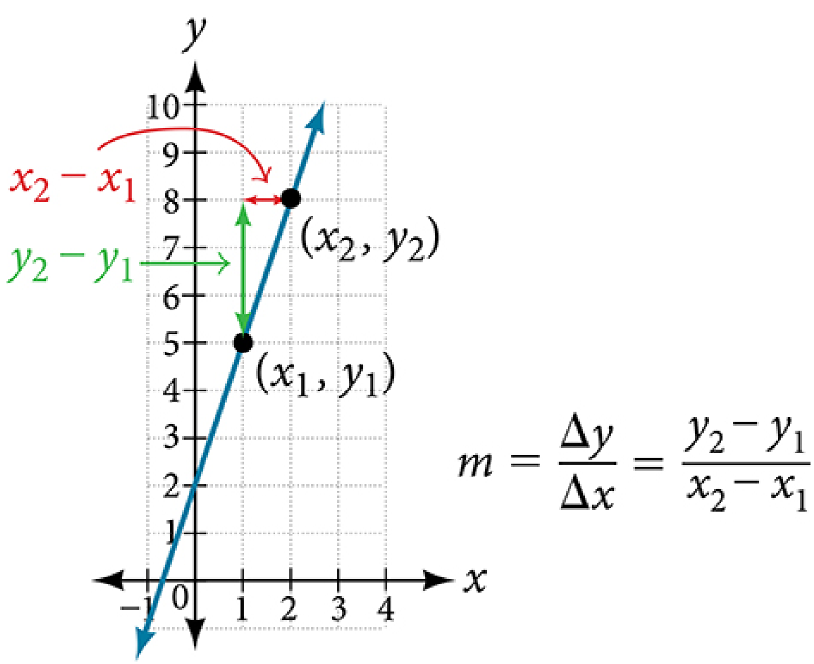



Figure 6.13: The slope of a function is calculated by the change in $y$ divided by the change in $x$. It does not matter which coordinate is used as the $(x_2, y_2)$ and which is the $(x_1, y_1)$, as long as each calculation is started with the elements from the same coordinate pair.

Figure 6.13 indicates how the slope of the line between the points, $(x_1, y_1)$ and $(x_2, y_2)$, is calculated. Recall that the slope measures steepness or slant. The greater the absolute value of the slope, the steeper the slant is.

**Definition 6.2.2** Given two points of a linear function, calculate and interpret the slope:

- Determine the units for output and input values.
- Calculate the change of output values and change of input values.
- Interpret the slope as the change in output values per unit of the input value.

### Determining Whether a Linear Function Is Increasing, Decreasing, or Constant

A linear function may be increasing, decreasing, or constant. For an increasing function, as with the train example, the output values increase as the input values increase. The graph of a growing function has a positive slope. A line with a positive slope slants upward from left to right. For a decreasing function, the slope is negative. The output values decrease as the input values increase. A line with a negative slope slants downward from left to right. If the function is constant, the output values are the same for all input values, so the slope is zero. A line with a slope of zero is horizontal.

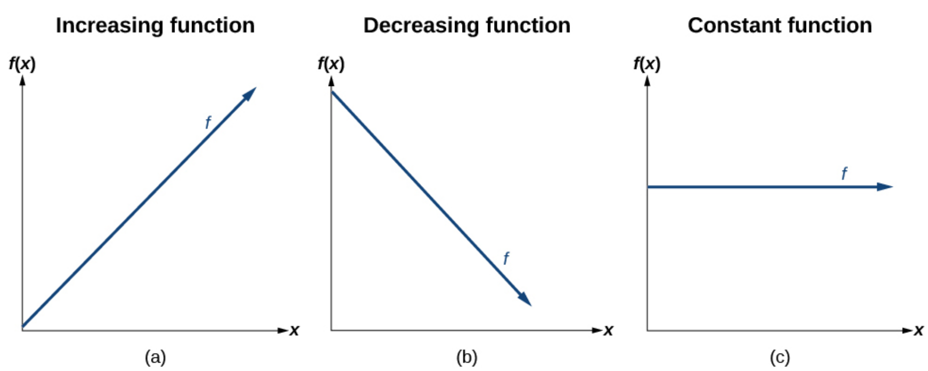



Figure 6.14: The monotony of linear functions

**Definition 6.2.3** The slope determines if the function is an increasing linear function, a decreasing linear function, or a constant function.

- $f(x) = mx + b$ is an increasing function if $m > 0$
- $f(x) = mx + b$ is an decreasing function if $m < 0$
- $f(x) = mx + b$ is a constant function if $m = 0$

■ **Example 6.19 — Finding the Slope of a Linear Function.** If $f(x)$ is a linear function, and $(3,-2)$ and $(8,1)$ are points on the line, find the slope. Is this function increasing or decreasing?

**Solution:** The coordinate pairs are $(3,-2)$ and $(8,1)$. To find the rate of change, we divide the change in output by the change in input.

$$m = \frac{\text{change in output (rise)}}{\text{change in input (run)}} = \frac{1-(-2)}{8-3}$$

We could also write the slope as $m = 0.6$. The function is increasing because $m > 0$.

■

■ **Example 6.20 — Deciding Whether a Function Is Increasing, Decreasing, or Constant.** Some recent studies suggest that a teenager sends an average of 60 texts per day. For each of the following scenarios, find the linear function describing the relationship between the input and output values. Then, determine whether the function's graph is increasing, decreasing, or constant.

- The total number of texts a teen sends is considered a function of time in days. The input is the number of days, and the output is the total number of texts sent.
- A teen has a limit of 500 texts per month in their data plan. The input is the number of days, and the output is the total number of texts remaining for the month.
- A teen has an unlimited number of texts in their data plan for \$50 per month. The input is the number of days, and the output is the total cost of texting each month.

**Solution:**

- We can represent the function as $f(x) = 60x$ where $x$ is the number of days. The slope, 60, is positive, so the function is increasing; this makes sense because the total number of texts increases with each day.
- We can represent the function as $f(x) = 500 - 60x$ where $x$ is the number of days. In this case, the slope is negative, so the function is decreasing; this makes sense because the number of texts remaining decreases each day, and this function represents the number of texts remaining in the data plan after x days.
- We can describe the function as $f(x) = 50$ because the number of days does not affect the total cost. The slope is 0, so the function is constant.

■

■ **Example 6.21 — Finding the Population Change from a Linear Function.** A city's population increased from 23,400 to 27,800 between 2008 and 2012. Find the change in population per year if we assume the change was constant from 2008 to 2012.

**Solution:** The population increased by $27,800 - 23,400 = 4400$ people over the four-year time interval. The rate of change relates the change in population to the change in time. To find the rate of change, divide the difference among the number of people by the number of years.

$$m = \frac{\text{change in output (rise)}}{\text{change in input (run)}} = \frac{4,400\text{people}}{4\text{years}} = 1,100\frac{\text{people}}{\text{years}}$$

So the population increased by 1,100 people per year.

■

### 6.2.3 Modeling with Linear Functions

#### Building Linear Models from Verbal Descriptions

When building linear models to solve problems involving quantities with a constant rate of change, we typically follow the same problem strategies we would use for any function. Let's briefly review them:

- Identify changing quantities, and then define descriptive variables to represent those quantities. When appropriate, sketch a picture or define a coordinate system.
- Carefully read the problem to identify important information. Look for information that provides values for the variables or parts of the functional model, such as slope and initial value.
- Carefully read the problem to determine what we are trying to find, identify, solve, or interpret.
- Identify a solution pathway from the provided information to what we are trying to find. Often this will involve checking and tracking units, building a table, or even finding a formula for the function used to model the problem.
- When needed, write a formula for the function.
- Solve or evaluate the function using the formula.
- Reflect on whether your answer is reasonable for the given situation and whether it makes sense mathematically.
- Show your result using appropriate units and answer in complete sentences when necessary.

■ **Example 6.22** Emily is a college student who plans to spend a summer in Seattle. She has saved $\$3,500$ for her trip and anticipates spending $\$400$ each week on rent, food, and activities.

- How can we write a linear model to represent her situation?.
- What would be the x-intercept, and what can she learn from it?.

We can create a linear function model to answer these and related questions. Models such as this can be beneficial for analyzing relationships and making predictions based on those relationships. In the student in Seattle situation, there are two changing quantities: time and money. The amount of money she has while on vacation depends on how long she stays. We can use this information to define our variables, including units.

First, we define the input and output of the linear function.

**Output: M, money remaining, in dollars**
**Input, t, time, in weeks**

We can also identify the initial value and the rate of change.

**Initial Value: She saved $\$3,500$, so $\$3,500$ is the initial value for $M$.**
**Rate of Change: Spending $\$400$ each week, so $\$400$ per week is the rate of change, or slope.**

Notice that the unit of dollars per week matches the unit of our output variable divided by our input variable. Also, because the slope is negative, the linear function is decreasing; this should make sense because she is spending money each week.

The rate of change is constant, so that we can start with the linear model $M(t) = mt + b$. Then we can substitute the intercept and slope provided.

$$M(t) = \underbrace{m}_{-400}t + \underbrace{b}_{3500}$$

$$M(t) = -400t + 3500$$

Figure 6.15: M as linear function of t

To find the t-intercept (horizontal axis intercept), we set the output to zero and solve for the input. $0 = -400t + 3500$ then $t = 3500/400 = 8.75$.

The $t$-intercept (horizontal axis intercept) is 8.75 weeks. Because this represents the input value when the output will be zero, we could say that Emily will have no money left after 8.75 weeks.

When modeling any real-life scenario with functions, there is typically a limited domain over which that model will be valid—almost no trend continues indefinitely. Here the domain refers to the number of weeks. Talking about input values less than zero doesn't make sense in this case. A negative input value could refer to several weeks before she saved $\$3,500$, but the scenario discussed poses the question once she saved $\$3,500$ because this is when her trip and subsequent spending start. It is also likely that this model is not valid after the $t$-intercept (horizontal axis intercept) unless Emily uses a credit card and goes into debt. The domain represents the set of input values, so the likely domain for this function is $0 \leq t \leq 8.75$.

In this example, we were given a written description of the situation. We followed the steps of modeling a problem to analyze the information. However, the information provided may not always be the same. Sometimes it might provide us with an intercept. Other times supply an output value. We must carefully analyze the information we are given and use it appropriately to build a linear model.

■

**Proposition 6.2.1** Given a word problem that includes two pairs of input and output values, use the linear function to solve a problem.

(1) Identify the input and output values.
(2) Convert the data to two coordinate pairs.
(3) Find the slope.
(4) Write the linear model.
(5) Use the model to make a prediction by evaluating the function at a given x-value.
(6) Use the model to identify an x-value that results in a given y-value.
(7) Answer the question posed.

■ **Example 6.23 — Using a Linear Model to Investigate a Town's Population.** A town's population has been growing linearly. In 2004, the population was 6,200. By 2009, the population had grown to 8,100. Assume this trend continues.

(1) Predict the population in 2013.

(2) Identify the year the population will reach 15,000.

**Solution:**

The two changing quantities are the population size and time. While we could use the actual year value as the input quantity, doing so tends to lead to very cumbersome equations because the y-intercept would correspond to the year 0, more than 2000 years ago!

To make computation a little nicer, we will define our input as the years since 2004.

**Input: t, years since 2004**
**Output: P(t), the town's population**

We would first need an equation for the population to predict the population in 2013 ( t=9 ). Likewise, to find when the people would reach 15,000, we would need to solve for the input that would provide an output of 15,000. To write an equation, we need the initial value, the rate of change, or slope.

To determine the rate of change, we will use the change in output per change in input.

$$m = \frac{\textbf{change in input}}{\textbf{change in output}}$$

The problem gives us two input-output pairs. To match our defined variables, the year 2004 would correspond to $t = 0$, giving the point $(0, 6200)$. Notice that through our clever choice of variable definition, we have "given" ourselves the $y$-intercept of the function. The year 2009 would correspond to t=5, giving the point $(5, 8100)$.

The two coordinate pairs are $(0, 6200)$and $(5, 8100)$. Recall that we encountered examples that provide us with two points earlier in the chapter. We can use these values to calculate the slope.

$$\begin{aligned} m &= \frac{8100 - 6200}{5 - 0} \\ &= \frac{1000}{5} \\ &= 380 \text{ people per year.} \end{aligned}$$

We already know the y-intercept of the line, so we can immediately write the equation:

$$P(t) = 380t + 6200$$

To predict the population in 2013, we evaluate our function at $t = 9$.

$$\begin{aligned} P(9) &= 380(9) + 6200 \\ &= 9620 \end{aligned}$$

If the trend continues, our model predicts a population of 9,620 in 2013.

To find when the population will reach 15,000, we can set $P(t) = 15000$ and solve for $t$.

$$\begin{aligned} 15000 &= 380t + 6200 \\ 8800 &= 380t \\ t &\approx 23.158 \end{aligned}$$

Our model predicts the population will reach 15,000 in a little more than 23 years after 2004, or around 2027. ■

### Using a Diagram to Build a Model

It is helpful for many real-world applications to draw a picture to understand how we may use the variables representing the input and output to answer a question. To draw the picture, consider what the problem is asking for. Then, determine the input and the output. The diagram should relate to the variables. Often, geometrical shapes or figures are drawn. Distances are often traced out. If a right triangle is sketched, the Pythagorean Theorem relates the sides. If a rectangle is outlined, labeling width and height is helpful.

■ **Example 6.24 — Using a Diagram to Model Distance Walked.** Anna and Emanuel start at the same intersection. Anna walks east at 4 miles per hour, while Emanuel walks south at 3 miles per hour. They are communicating with a two-way radio that has a range of 2 miles. How long after they start walking will they fall out of radio contact?

**Solution:**

In essence, we can partially answer this question by saying they will fall out of radio contact when they are 2 miles apart, which leads us to ask a new question:

"How long will it take them to be 2 miles apart"?

In this problem, our changing quantities are time and position, but ultimately we need to know how long will it take for them to be 2 miles apart. We can see that time will be our input variable, so we'll define our input and output variables.

**Input: t, time in hours**
**Output: A(t), distance in miles, and E(t), distance in miles**

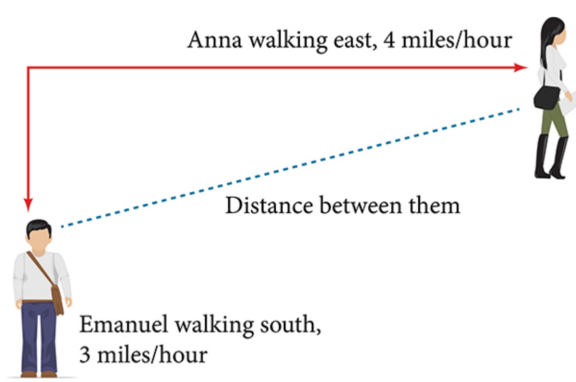



Figure 6.16: Diagram of the distance

Because it is not obvious how to define our output variable, we'll start by drawing a picture such as the figure above 6.16.

Initial Value: They both start at the same intersection, so when t=0, the distance traveled by each person should also be 0. Thus the initial value for each is 0.

Rate of Change: Anna is walking 4 miles per hour, and Emanuel is walking 3 miles per hour, which are both rates of change. The slope for $A$ is four, and the slope for $E$ is 3.

Using those values, we can write formulas for the distance each person has walked.

$$A(t) = 4t$$
$$E(t) = 3t$$

For this problem, the distances from the starting point are essential. To notate these, we can define a coordinate system, identifying the "starting point" at the intersection where they both started. Then we can use the variable $A$, which we introduced above, to represent Anna's position, and define it as a measurement from the starting point in the eastward direction. Likewise, we can use the variable $E$ to represent Emanuel's position, measured from the starting point in the southward direction. Note that in defining the coordinate system, we specified both the starting point of the measurement and the direction of measure.

We can then define a third variable, $D$, to measure the distance between Anna and Emanuel. Showing the variables on the diagram is often helpful, as we can see from Figure 6.17.

Recall that we need to know how long it takes for $D$, the distance between them, to equal 2 miles. Notice that for any given input $t$, the outputs $A(t)$,$E(t)$, and $D(t)$ represent distances.

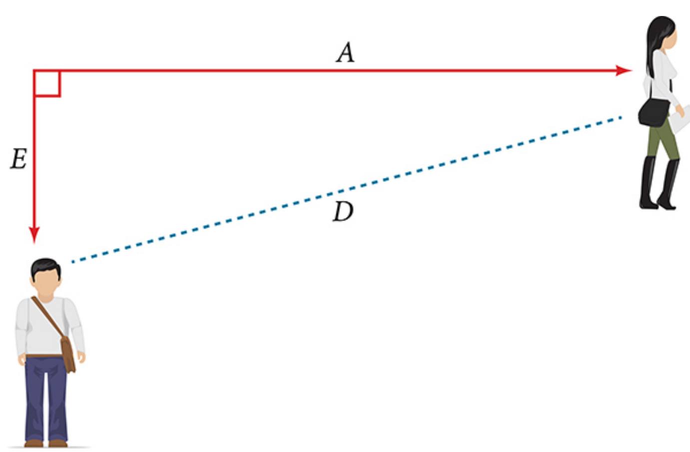



Figure 6.17: Diagram of the distance

Figure 6.17 shows us that we can use the Pythagorean Theorem because we have drawn a right angle. Using the Pythagorean Theorem, we get:

$$\begin{aligned}
D(t)^2 &= A(t)^2 + E(t)^2 \\
&= (4t)^2 + (3t)^2 \\
&= 16t^2 + 9t^2 \\
&= 25t^2 \\
D(t) &= \pm\sqrt{25t^2} = \pm 5|t|.
\end{aligned}$$

In this scenario, we are considering only positive values of t, so our distance $D(t)$ will always be positive. We can simplify this answer to D(t)=5t; this means that the distance between Anna and Emanuel is also a linear function. Because $D$ is a linear function, we can now answer the question of when the distance between them will reach 2 miles. We will set the output $D(t) = 2$ and solve for $t$.

$$\begin{aligned}
D(t) &= 2 \\
5t &= 2 \\
t &= 2/5 = 0.4
\end{aligned}$$

They will fall out of radio contact in 0.4 hours or 24 minutes.

■

■ **Example 6.25 — Using a Diagram to Model Distance Between Cities.** There is a straight road leading from the town of Westborough to Agritown, 30 miles east and 10 miles north. Partway down this road, it junctions with a second road, perpendicular to the first, leading to the town of Eastborough. If the town of Eastborough is located 20 miles directly east of the city of Westborough, how far is the road junction from Westborough?

**Solution:**

It might help here to draw a picture of the situation. See Figure 6.18.

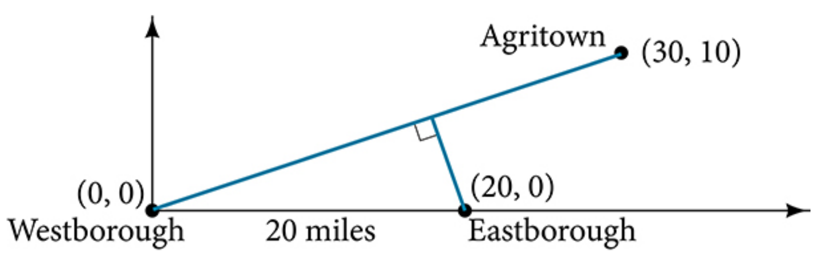



Figure 6.18: Diagram of the distance

It would then be helpful to introduce a coordinate system. While we could place the origin anywhere, placing it at Westborough seems convenient. This puts Agritown at coordinates (30,10), and Eastborough at $(20,0)$.

Using this point along with the origin, we can find the slope of the line from Westborough to Agritown.

$$m = \frac{10-0}{30-0} = \frac{1}{3}$$

Now we can write an equation to describe the road from Westborough to Agritown.

$$W(x) = \frac{1}{3}x$$

From this, we can determine the perpendicular road to Eastborough will have a slope of $m = -3$. Because the town of Eastborough is at the point $(20,0)$, we can find the equation.

$$\begin{aligned} E(x) &= -3x+b \\ 0 &= -3(20)+b \quad \text{substitute}(20,0) \text{ into the equation} \\ b &= 60 \\ E(x) &= -3x+60 \end{aligned}$$

We can now find the coordinates of the junction of the roads by finding the intersection of these lines. Setting them equal,

$$\begin{aligned} \frac{1}{3}x &= -3x+60 \\ \frac{10}{3}x &= 60 \\ 10x &= 180 \\ y &= W(18) = \frac{1}{3}(18) = 6 \end{aligned}$$

The roads intersect at the point $(18,6)$. We can now find the distance from Westborough to the junction using the distance formula.

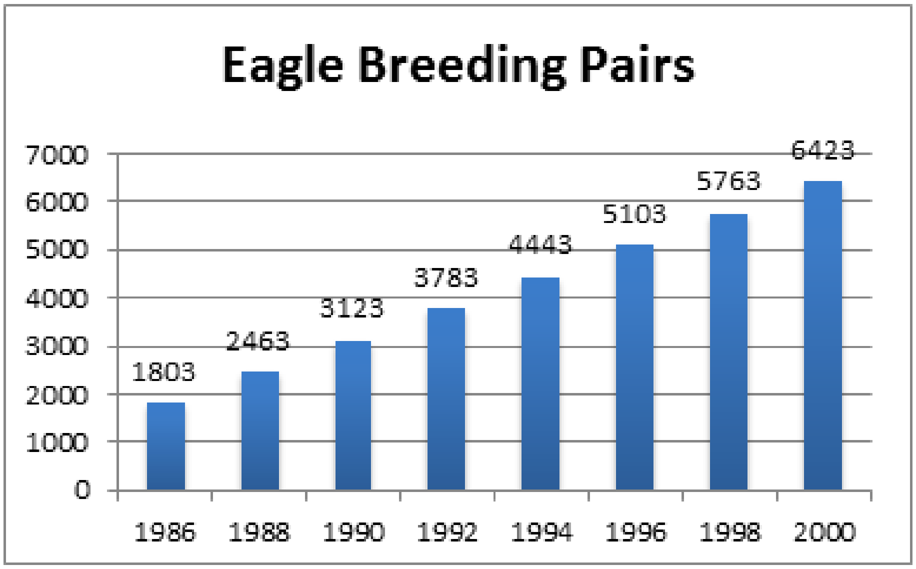
$$\begin{aligned} \text{distance} &= \sqrt{(x_2-x_1)^2+(y_2-y_1)^2} \\ &= \sqrt{(18-0)^2+(6-0)^2} \\ &\approx 18.974 \text{ miles} \end{aligned}$$

∎

One excellent use of linear models is to take advantage of the fact that the graphs of these functions are lines; this means real-world applications discussing maps need linear functions to model the distances between reference points.

### 6.2.4 Exercises 6.2

**Problem 6.23** From the early 1980s through 2000, states kept records of the number of bald eagle breeding pairs. Here is a bar graph showing the approximate number of pairs of breeding eagles from 1986 to 2000.




1. Use the data in the bar graph to complete the table. In the table, let $t$ represent the number of years after 1986, and let $P(t)$ be the number of breeding pairs of eagles.

| $t$ | 0 | | 4 | | | 10 | | 14 |
|---|---|---|---|---|---|---|---|---|
| $P(t)$ | 1803 | | 3123 | | | 5103 | | 6423 |

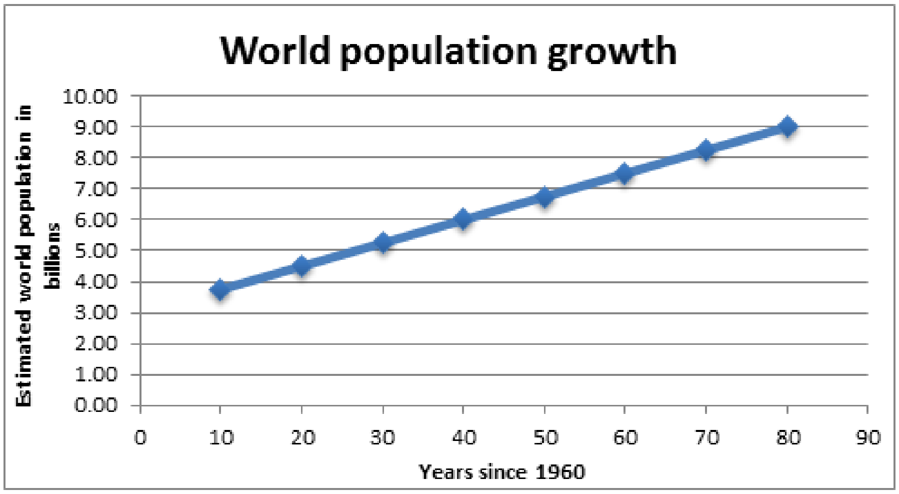
2. Decide whether the number of breeding pairs seems to be growing linearly. Write your conclusion in a complete sentence.

3. Find the growth rate. Express your answer using correct units.

4. Find a formula for the function $P(t)$ expressing the number of breeding pairs of eagles in terms of $t$, the number of years after 1986.

5. In 2006, the year before bald eagles were removed from the Endangered Species List. The number of breeding pairs was 9,789. Does your linear model over-or under-estimate the actual number of breeding pairs? Explain your reasoning in a complete sentence.

**Problem 6.24** The United States Census uses world population information to model population growth. The table and graph below show data for one linear model that the Census Bureau created.




This is also shown in the following table:

| Years Since 1960 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|---|---|
| Worlds Population (in billions) | 3.75 | 4.50 | 5.25 | 6.00 | 6.75 | 7.50 | 8.25 | 9.0 |

1. Does the graph show information about a vertical intercept?
2. Pick two ordered pairs corresponding to points on the graph. Use those ordered pairs to calculate the rate of change. Make sure you write the correct units in your answer.
3. Compare the rate of change you calculated with the answers other people found. Are all of the rates of change the same? If so, why?
4. Let $t$ represent the number of years since 1960 and let $P(t)$ represent the corresponding world population, in billions. Use the steps to create a linear function from two data points to find an equation that expresses the linear relationship between $P(t)$ and $t$.

**Problem 6.25**

Explain how to find the input variable in a word problem that uses a linear function.

**Problem 6.26**

Explain how to find the output variable in a word problem that uses a linear function.

**Problem 6.27**

Explain how to interpret the initial value in a word problem that uses a linear function.

**Problem 6.28**

Explain how to determine the slope in a word problem that uses a linear function.

**Problem 6.29** For the following exercises, consider this scenario:

A town's population has been decreasing at a constant rate. In 2010 the population was 5,900. By 2012 the population had dropped 4,700. Assume this trend continues.

- Predict the population in 2016.
- Identify the year in which the population will reach 0.

**Problem 6.30** For the following exercises, consider this scenario: A town's population has increased at a constant rate. In 2010 the population was 46,020. By 2012 the population had risen to 52,070. Assume this trend continues.

- Predict the population in 2016.
- Identify the year in which the population will reach 75000.

**Problem 6.31** For the following exercises, consider this scenario: A town has an initial population of 75,000. It grows at a constant rate of 2,500 per year for five years.

- Find the linear function that models the town's population $P$ as a function of the year, $t$, where $t$ is the number of years since the model began.
- Find a reasonable domain and range for the function $P$.
- If the function P is graphed, find and interpret the x- and y-intercepts.
- If the function P is graphed, find and interpret the slope of the function.
- When will the population reach 100,000?
- What is the population in the year 12 years from the onset of the model?

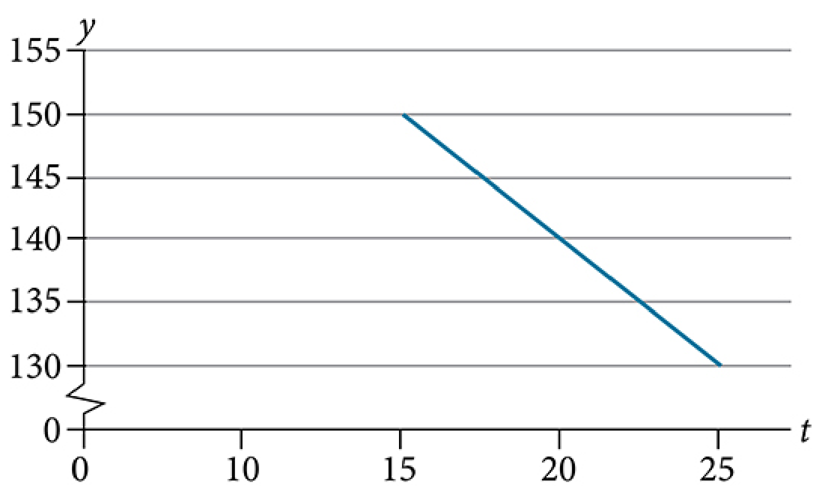
**Problem 6.32** For the following exercises, consider this scenario: A town's population has increased at a constant rate. In 2012 the population was 56,000. By 2015 the population had increased to 82,000. Assume this trend continues. What was the population in 2016?

**Problem 6.33** For the following exercises, consider this scenario: The weight of a newborn is 7.5 pounds. The baby gained one-half pound a month for its first year.

- Find the linear function that models the baby's weight $W$ as a function of the age of the baby, in months, $t$.
- Find a reasonable domain and range for the function $W$.
- If the function $W$ is graphed, find and interpret the $x$- and $y$-intercepts.
- If the function $W$ is graphed, find and interpret the slope of the function.
- When did the baby weight 10.4 pounds?
- What is the output when the input is 6.2?

**Problem 6.34**

For the following exercises, use the graph in the Figure below , which shows the profit, $y$, in thousands of dollars, of a company in a given year, $t$, where $t$ represents the number of years since 1980.

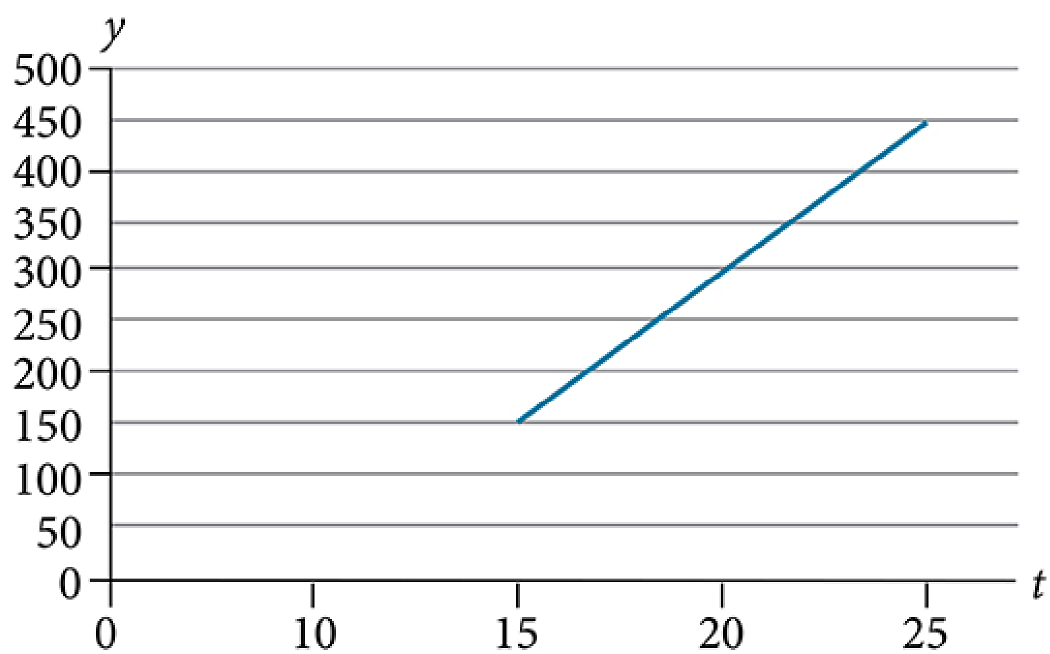



- Find the linear function y, where y depends on t, the number of years since 1980.
- Find and interpret the y-intercept.
- Find and interpret the x-intercept.
- Find and interpret the slope.

**Problem 6.35** For the following exercises, consider this scenario: The number of people afflicted with the common cold in the winter months steadily decreased by 205 each year from 2005 until 2010. In 2005, 12,025 people were afflicted.

- Find the linear function that models the number of people inflicted with the common cold $C$ as a function of the year, $t$.
- When will the output reach 0?
- In what year will the number of people be 9,700?

**Problem 6.36** For the following exercises, use the graph in the Figure below, which shows the profit, y, in thousands of dollars, of a company in a given year, t, where t represents the number of years since 1980.




- Find the linear function y, where y depends on t, the number of years since 1980.
- Find and interpret the x-intercept and y-intercept.
- Find and interpret the slope.
- When will it reach 600?

### Problem 6.37

For the following exercises, use the median home values in Mississippi and Hawaii (adjusted for inflation) shown in the table below. Assume that the house values are changing linearly.

| Year | Mississippi | Hawaii |
|---|---|---|
| 1950 | \$25,200 | \$74,400 |
| 2000 | \$71,400 | \$272,700 |

- In which state has home values increased at a higher rate?
- If these trends were to continue, what would be the median home value in Mississippi in 2010?
- If we assume the linear trend existed before 1950 and continues after 2000, the two states' median house values will be (or were) equal in what year? (The answer might be absurd.)

**Problem 6.38** For the following exercises, use the median home values in Indiana and Alabama (adjusted for inflation) shown in the table below. Assume that the house values are changing linearly.

| Year | Indiana | Alabama |
|---|---|---|
| 1950 | \$37,700 | \$27,100 |
| 2000 | \$94,300 | \$85,100 |

- In which state have home values increased at a higher rate?
- If these trends were to continue, what would be the median home value in Indiana in 2010?
- If we assume the linear trend existed before 1950 and continues after 2000, the two states' median house values will be (or were) equal in what year? (The answer might be absurd.)

**Problem 6.39** In 2004, the school population was 1001. By 2008 the population had grown to 1697. Assume the population is changing linearly.

- How much did the population grow between the years 2004 and 2008?
- How long did it take the population to grow from 1001 students to 1697 students?
- What is the average population growth per year?
- What was the population in the year 2000?
- Find an equation for the population, P, of the school t years after 2000.
- Using your equation, predict the school's population in 2011.

**Problem 6.40** When hired at a new job selling jewelry, you are given two payment options:

- Option A: Base salary of $\$17,000$ a year with commission of $12\%$ of your sales.
- Option B: Base salary of $\$20,000$ a year with a commission of $5\%$ of your sales.

How much jewelry would you need to sell for option A to produce a more considerable income?

**Problem 6.41** A phone company has a monthly cellular plan where a customer pays a flat monthly fee and then a certain amount of money per minute used on the phone. If a customer uses 410 minutes, the monthly cost will be $71.50. If the customer uses 720 minutes, the monthly fee will be $118.

- Find a linear equation for the monthly cost of the cell plan as a function of x, the number of monthly minutes used. Interpret the slope and y-intercept of the equation.
- Use your equation to find the total monthly cost if 687 minutes are used.

**Problem 6.42** In 1991, the moose population in a park was measured to be 4,360. By 1999, the population was observed to be 5,880. Assume the population continues to change linearly.

- Find a formula for the moose population, P since 1990.
- What does your model predict the moose population to be in 2003?

**Problem 6.43** In 2003, the owl population in a park was measured to be 340. By 2007, it was estimated again to be 285. The population changes linearly starting from 2003.

- Find a formula for the owl population, P. Let the input be years since 2003.
- What does your model predict the owl population to be in 2012?

**Problem 6.44** The Federal Helium Reserve held about 16 billion cubic feet of helium in 2010 and is being depleted by about 2.1 billion cubic feet each year.

- Give a linear equation for the remaining federal helium reserves, $R$, in terms of $t$, the number of years since 2010.
- In 2015, what will the helium reserves be?
- If the depletion rate doesn't change, in what year will the Federal Helium Reserve be depleted?

**Problem 6.45**

When hired at a new job selling electronics, you are given two payment options:

- Option A: Base salary of $\$20,000$ a year with commission of $12\%$ of your sales.
- Option B: Base salary of $\$26,000$ a year with a commission of $3\%$ of your sales.

How much electronics would you need to sell for option A to produce a larger income?

## 6.3 Exponential Modeling

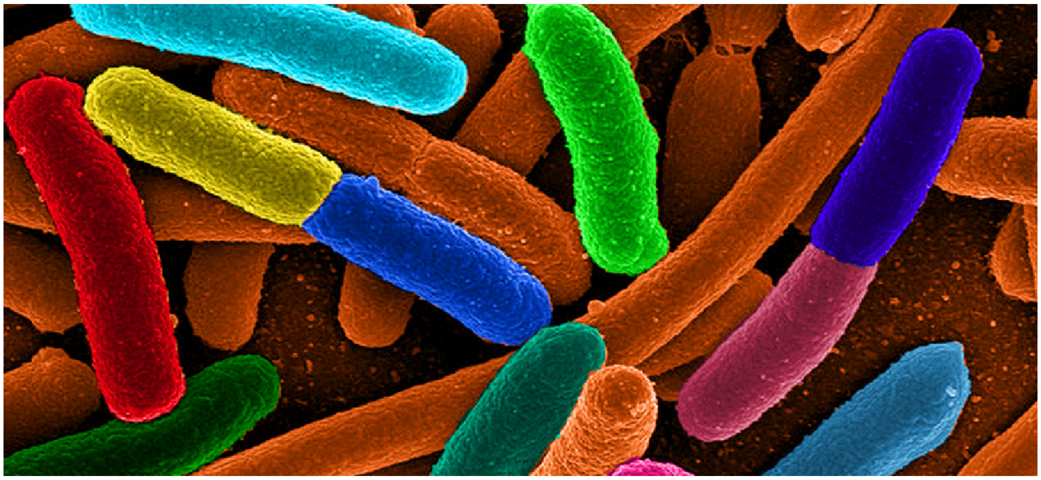


Figure 6.19: Electron micrograph of E.Coli bacteria (credit: "Mattosaurus," Wikimedia)

Focus on a square centimeter of your skin. Look closer. Closer still. If you could look closely enough, you would see hundreds of thousands of microscopic organisms. They are bacteria, not only on your skin but in your mouth, nose, and even your intestines. The bacterial cells in your body at any given moment outnumber your cells. But that is no reason to feel bad about yourself. While some bacteria can cause illness, many are healthy and essential to the body.

Bacteria commonly reproduce through binary fission, during which one bacterial cell splits into two. When conditions are right, bacteria can reproduce very quickly. Unlike humans and other complex organisms, the time required to form a new generation of bacteria is often a matter of minutes or hours, as opposed to days or years.

For simplicity's sake, suppose we begin with a culture of one bacterial cell that can divide every hour. Table 6.5 shows the number of bacterial cells at the end of each subsequent hour. The single bacterial cell leads to over one thousand bacterial cells in just ten hours! And if we were to extrapolate the table to twenty-four hours, we would have over 16 million!

| Hour | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bacteria | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 | 1024 |

Table 6.5: Bacteria Growth

In this section, we will explore exponential functions, which can be used for, among other things, modeling growth patterns such as those found in bacteria. This function has numerous real-world applications for modeling and interpreting data.

India is the second-most populous country in the world, with a population of about 1.25 billion people in 2013. The population is growing at a rate of about 1.2% each year. If this rate continues, the population of India will exceed China's population by the year 2031. When populations grow rapidly, we often say that the growth is "exponential," meaning something is growing quickly. However, to a mathematician, the term exponential growth has a specific meaning. In this section, we will look at exponential functions, which model this kind of rapid growth.

## 6.3.1 Identifying Exponential Functions

When exploring linear growth, we observed a constant rate of change—a constant number by which the output increased for each unit increase in input. For example, in the equation f(x)=3x+4, the slope tells us the output increases by three each time the input increases by 1. The scenario in the India population example is different because we have a percent change per unit time (rather than a constant change) in the number of people.

### Defining an Exponential Function

A study found that the percentage of the population who are vegans in the United States doubled from 2009 to 2011. In 2011, 2.5% of the population was vegan, adhering to a diet that does not include animal products—no meat, poultry, fish, dairy, or eggs. If this rate continues, vegans will make up 10% of the U.S. population in 2015, 40% in 2019, and 80% in 2021.

What exactly does it mean to grow exponentially? What does the word double have in common with a percent increase? People toss these words around errantly. Are these words used correctly? The words indeed appear frequently in the media.

Percent change refers to a change based on a percent of the original amount. Exponential growth refers to an increase based on a constant multiplicative rate of change over equal increments of time, that is, a percent increase of the original amount over time. Exponential decay refers to a decrease based on a constant multiplicative rate of change over equal increments of time, that is, a percent decrease of the original amount over time.

For us to gain a clear understanding of exponential growth, let us contrast exponential growth with linear growth. We will construct two functions. The first function is exponential. We will start with an input of 0 and increase each input by 1. We will double the corresponding consecutive outputs. The second function is linear. We will start with an input of 0 and increase each input by 1. We will add 2 to the corresponding consecutive outputs. See Table 6.6.

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $f(x) = 2^x$ | 1 | 2 | 4 | 8 | 16 | 32 | 64 |
| $g(x) = 2x$ | 0 | 2 | 4 | 6 | 8 | 10 | 12 |

Table 6.6: Exponential vs Linear growth

From Table 6.6 we can infer that for these two functions, exponential dwarfs linear growth.

- **Exponential growth** refers to the original value from the range increases by the same percentage over equal increments found in the domain.

$$\frac{\text{New Value} - \text{Old Value}}{\text{Old Value}} = \text{Constant } \%$$

- **Linear growth refers** to the original value from the range increases by the same amount over equal increments found in the domain.

$$\text{New Value} - \text{Old Value} = \text{Constant}$$

The difference between "the same percentage" and "the same amount" is quite significant. For exponential growth, over equal increments, the constant multiplicative rate of change resulted in doubling the output whenever the input increased by one. For linear growth, the continual additive rate of change over equal increments resulted in adding 2 to the output whenever the intake was raised by one.

The general form of the exponential function is $f(x) = ab^x$, where $a$ is any nonzero number, $b \neq 1$ is a positive real number.

- If $b > 1$, the function grows at a rate proportional to its size.
- If $0 < b < 1$, the function decays at a rate proportional to its size.

Let's look at the function $f(x) = 2x$ from our example. We will create a new table (Table 6.7) to determine the domain's corresponding outputs over an interval from $-3$ to 3.

| x | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| $f(x) = 2^x$ | $2^{-3} = \frac{1}{8}$ | $2^{-2} = \frac{1}{4}$ | $2^{-1} = \frac{1}{2}$ | $2^0 = 1$ | $2^1 = 2$ | $2^2 = 4$ | $2^3 = 8$ |

Table 6.7: Exponential vs Linear growth

Let us examine the graph of $f$ by plotting the ordered pairs we observe in the Table in Figure 6.19, and then make a few observations.

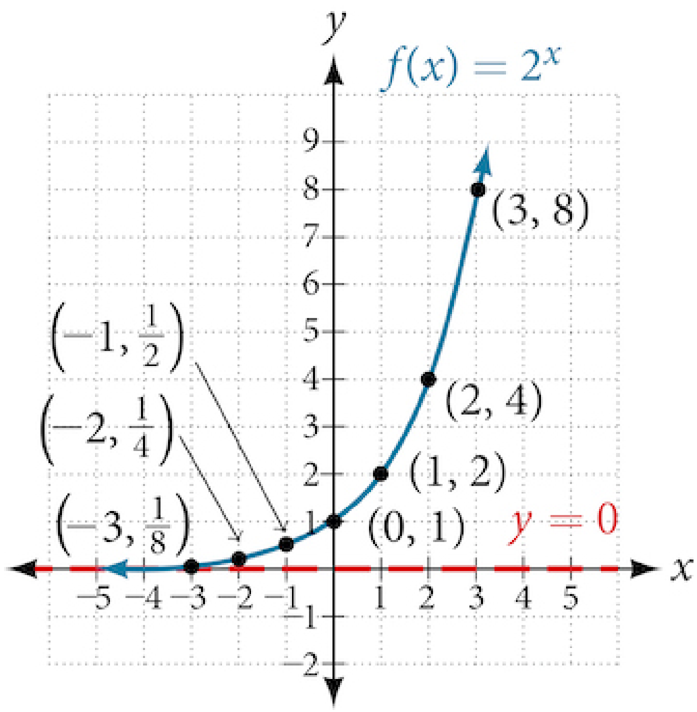



Figure 6.20: The $2^x$ graph

Let's define the behavior of the graph of the exponential function $f(x) = 2^x$ and highlight some of its essential characteristics:

- the domain is $(-\infty, \infty)$,
- the range is $(0, \infty)$,
- as $x \to \infty$, $f(x) \to \infty$,
- as $x \to -\infty$, $f(x) \to 0$,
- $f(x)$ is always increasing,
- the graph of $f(x)$ will never touch the x-axis because base two raised to any exponent never has the result of zero.
- $y = 0$ is the horizontal asymptote.
- the $y$-intercept is 1.

**Definition 6.3.1 — Exponential function.** For any real number $x$, an exponential function is a function with the form.

$$f(x) = ab^x \tag{6.3}$$

where

- $a$ is a non-zero real number called the initial value and
- $b$ is any positive real number such that $b \neq 1$.
- The domain of $f$ is all real numbers.
- The range of $f$ is all positive real numbers if $a > 0$.
- The range of $f$ is all negative real numbers if $a < 0$.
- The $y$-intercept is $(0, a)$.

■ **Example 6.26** Which of the following equations are not exponential functions?

- $f(x) = 4^{3(x-2)}$
- $g(x) = x^3$
- $h(x) = (\frac{1}{3})^x$
- $j(x) = (-2)^x$
- $L(x) = 1095.6^x$

An exponential function has a constant as a base and an independent variable as an exponent. Thus, $g(x) = x^3$ does not represent an exponential function because the base is an independent variable. In fact, $g(x) = x^3$ is a power function.

Recall that the base $b$ of an exponential function is always a positive constant, and $b \neq 1$. Thus, $j(x) = (-2)^x$ does not represent an exponential function because the base, $-2$, is less than 0.

The number 1095.6 is a positive decimal, hence a positive real base. Therefore the function $L(x) = 1095.6^x$ can be classified as exponential function.

■

### 6.3.2 Evaluating Exponential Functions

Recall that the base of an exponential function must be a positive real number other than 1. Why do we limit the base $b$ to positive values? To ensure that the outputs will be real numbers. Observe what happens if the base is not positive:

- Let $b = -9$ and $x = \frac{1}{2}$. Then $f(x) = f(\frac{1}{2}) = (-9)^{\frac{1}{2}} = \sqrt{-9}$ which is not a real number.

Why do we limit the base to positive values other than 1? Because base one results in the constant function. Observe what happens if the base is one:

- Let $b = 1$. Then $f(x)(1)^x = 1$ for any value of $x$.

To evaluate an exponential function with the form $f(x) = b^x$, we simply substitute $x$ with the given value and calculate the resulting power.

■ **Example 6.27 — Evaluating $2^x$..** Let $f(x) = 2^x$. what is $f(3)$?

**Solution:**

$$f(x) = 2^x$$
$$f(3) = 2^3 \text{ substitute } x = 3$$
$$f(3) = 8 \text{ evaluate the power}$$

The order of operations is essential to evaluate an exponential function with a different basic form configuration. For instance:

Let $f(x) = 30(2)^x$. what is $f(3)$?

$$f(x) = 30(2)^x$$
$$f(3) = 30(2)^3 \text{ substitute } x = 3$$
$$f(3) = 30(8) \text{ simplify the power first}$$
$$f(3) = 240 \text{ multiply}$$

Note that if we don't follow the order of operations, the result would be incorrect as $f(3) = 30(2)^3 \neq 60^3 = 216000$.

■

■ **Example 6.28** Let $f(x) = 5(3)^{x+1}$. Evaluate $f(2)$ without using a calculator.

**Solution:**

$$f(x) = 5(3)^{x+1}$$
$$f(2) = 5(3)^{2+1} \text{ substitute } x = 3$$
$$f(2) = 5(3)^3 \text{ simplify the power first}$$
$$f(2) = 5(27) = 135 \text{ multiply}$$

■

### 6.3.3 Defining Exponential Growth

Because the output of exponential functions increases rapidly, the term "exponential growth" is often used in everyday language to describe anything that grows or expands quickly. However, we can define exponential growth more precisely in a mathematical sense. The function models exponential growth if the growth rate is proportional to the amount present.

**Definition 6.3.2** A function that models exponential growth grows by a rate proportional to the amount present. For any real number $x$ and any positive real numbers $a$ and $b$ such that $b \neq 1$, an exponential growth function has the form.

$$f(x) = ab^x$$

where

- $a$ is the initial or starting value of the function.
- $b$ is the growth factor or growth multiplier per unit $x$ .

In general terms, we have an exponential function, in which a constant base is raised to a variable exponent.

■ **Example 6.29 — differentiate between linear and exponential functions.** Let's consider two companies, A and B. Company A has 100 stores and expands by opening 50 new stores a year, so its growth can be represented by the function A(x)=100+50x. Company B has 100 stores and is expanding by increasing the number of stores by 50% each year, so its growth can be represented by the function $B(x) = 100(1+0.5)x$.

| Year,x | Stores, Company A | Stores, Company B |
|---|---|---|
| 0 | $100+50(0) = 100$ | $100(1+0.5)^0 = 100$ |
| 1 | $100+50(1) = 150$ | $100(1+0.5)^1 = 150$ |
| 2 | $100+50(2) = 200$ | $100(1+0.5)^2 = 225$ |
| 3 | $100+50(3) = 250$ | $100(1+0.5)^3 = 337.5$ |
| $x$ | $100+50(x)$ | $100(1+0.5)^x$ |

Table 6.8: A few years of growth for these companies are illustrated

The graphs comparing the number of stores for each company over five years are shown in Figure 6.21. We can see that, with exponential growth, the number of stores increases much more rapidly than with linear growth.

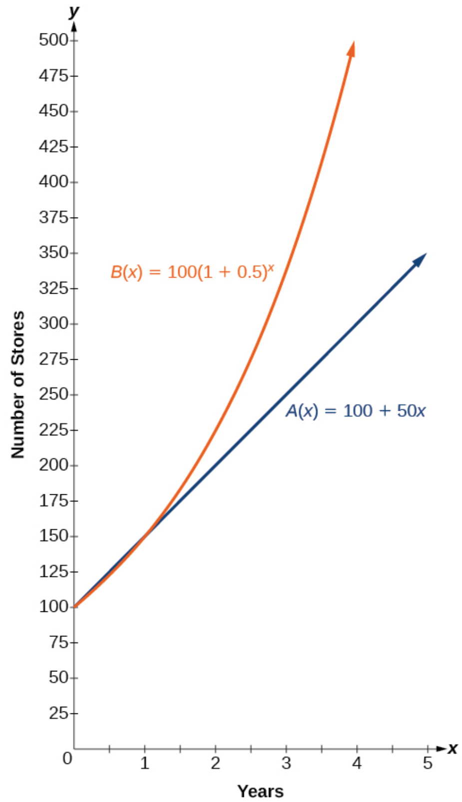



Figure 6.21: Company A vs Company B Growth.

Notice that the domain for both functions is $[0,\infty)$, and the range for both functions is $[100,\infty)$. After year 1, Company B always has more stores than Company A.

Now we will turn our attention to the function representing the number of stores for Company B, $B(x) = 100(1+0.5)x$. In this exponential function, 100 represents the initial number of stores, 0.50 represents the growth rate, and $1+0.5 = 1.5$ represents the growth factor. Generalizing further, we can write this function as $B(x) = 100(1.5)x$, where 100 is the initial value, 1.5 is the base, and $x$ is the exponent. ■

■ **Example 6.30** Evaluating a Real-World Exponential Model The population of India was about 1.25 billion in the year 2013, with an annual growth rate of about 1.2%. This situation is represented by the growth function $P(t) = 1.25(1.012)t$, where $t$ is the number of years since 2013. To the nearest thousandth, what will the population of India be in 2031?

**Solution:**

To estimate the population in 2031, we evaluate the models for$t = 18$, because 2031 is 18 years after 2013. Rounding to the nearest thousandth,

$$P(18) = 1.25(1.012)^{18} \approx 1.549$$

There will be about 1.549 billion people in India in the year 2031. ■

### 6.3.4 Introduction to the Exponential Function Graph

Exponential functions are used for many real-world applications such as finance, forensics, computer science, and most of the life sciences. Working with an equation that describes a real-world situation gives us a method for making predictions. Most of the time, however, the equation itself is not enough. We learn a lot about things by seeing their pictorial representations, which is precisely why graphing exponential equations is a powerful tool. It gives us another layer of insight for predicting future events. In this section, we introduce the graph of the exponential function and provide some basic facts that make exponential behavior easier to recognize.

#### Graphing Exponential Functions

Before we begin graphing, it is helpful to review exponential growth behavior. Recall the table of values for a function of the form $f(x) = b^x$ whose base is greater than one. We'll use the function $f(x) = 2^x$. Observe how the output values in Table 6.9 change as the input increases by 1.

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $f(x) = 2^x$ | 1 | 2 | 4 | 8 | 16 | 32 | 64 |

Table 6.9: Exponential growth of $2^x$

Each output value is the product of the previous output and the base, 2. We call base two the constant ratio. In fact, for any exponential function with the form $f(x) = ab^x$, b is the constant ratio of the function. If the input increases by 1, the output value will be the product of the base and the previous output, regardless of the value of $a$.

Notice from the table that:

- the output values are positive for all values of x;
- as x increases, the output values increase without bound; and
- as x decreases, the output values grow smaller, approaching zero.

Figure 6.22 shows the graph of $2^x$.

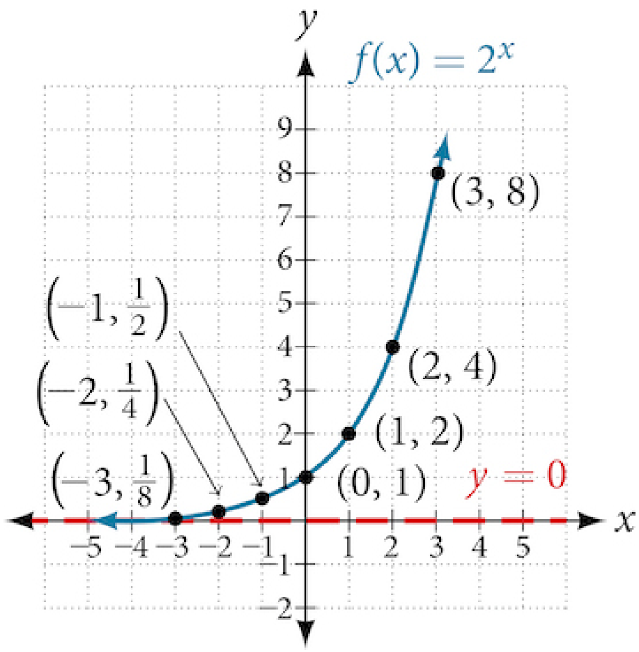



Figure 6.22: Exponential $2^x$ graph.

The domain of $f(x) = 2^x$ is all real numbers, the range is $(0, \infty)$, and the horizontal asymptote is $y = 0$.

To get a sense of the behavior of exponential decay, we can create a table of values for a function of the form $f(x) = b^x$ whose base is between zero and one. We'll use the function $g(x) = (\frac{1}{2})^x$. Observe how the output values in Table 6.10 change as input increases by 1.

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $g(x) = (\frac{1}{2})^x$ | 1 | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{16}$ | $\frac{1}{32}$ | $\frac{1}{64}$ |

Table 6.10: Exponential growth of $(\frac{1}{2})^x$

Again, because the input is increasing by 1, each output value is the product of the previous output and the base, or constant ratio $\frac{1}{2}$.

Notice from the table that:

- the output values are positive for all values of x;
- as x increases, the output values grow smaller, approaching zero; and
- as x decreases, the output values grow without bound.

Figure 6.23 shows the graph of $(\frac{1}{2})^x$.

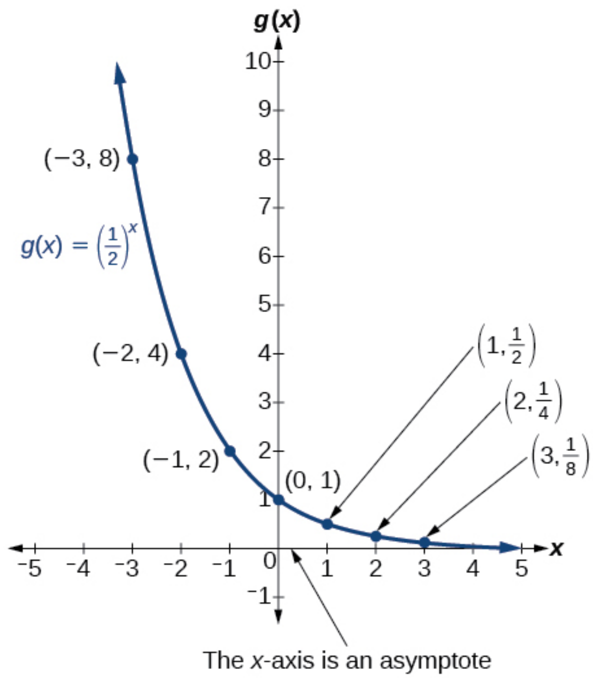



Figure 6.23: Exponential $(\frac{1}{2})^x$ graph.

The domain of $g(x) = (\frac{1}{2})^x$ is all real numbers, the range is $(0,\infty)$, and the horizontal asymptote is $y = 0$.

**Definition 6.3.3** An exponential function with the form $f(x) = b^x$, $b > 0$, $b \neq 1$, has these characteristics:

- one-to-one function
- domain: $(-\infty,\infty)$
- range: $(0,\infty)$
- $x$-intercept: none
- $y$-intercept: $(0,1)$
- increasing if $b > 1$
- decreasing if $b < 1$

Figure 6.24 compares the graphs of exponential growth and decay functions.

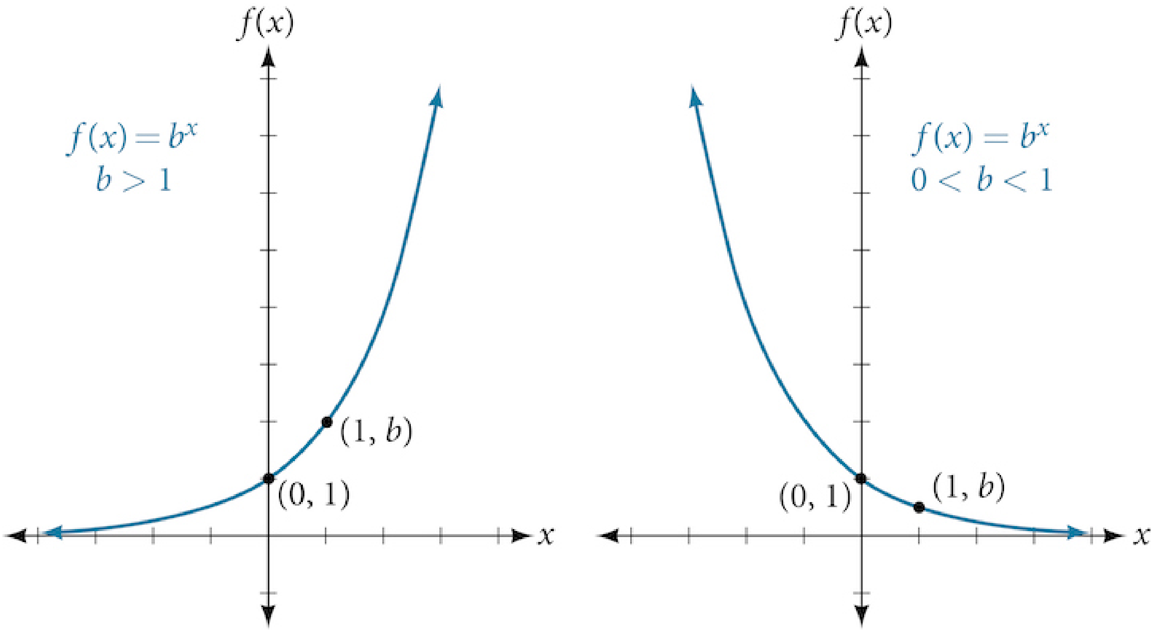



Figure 6.24: Exponential $b^x$ graph.

**Definition 6.3.4** Given an exponential function of the form $f(x) = b^x$, graph the function.

- Create a table of points.
- Plot at least 3 point from the table, including the $y$-intercept $(0,1)$.
- Draw a smooth curve through the points.
- State Domain and Range.

■ **Example 6.31** Sketch a graph of $f(x) = (0.25)^x$. State the domain, range, and asymptote.

**Solution:**

Before graphing, identify the behavior and create a table of points for the graph.

- Since $b = 0.25$ is between zero and one, and we know the function is decreasing. The left tail of the graph will increase without bound, and the right tail will approach the asymptote $y = 0$.
- Create a table of points as in Table 6.11.

| x | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| $f(x) = (0.25)^x$ | 64 | 16 | 4 | 1 | 0.25 | 0.06255 | 0.015625 |

Table 6.11: Exponential growth of $(0.25)^x$

- Plot the $y$-intercept, $(0,1)$, along with two other points. We can use $(-1,4)$ and $(1,0.25)$.

- Draw a smooth curve connecting the points as in Figure 6.24.

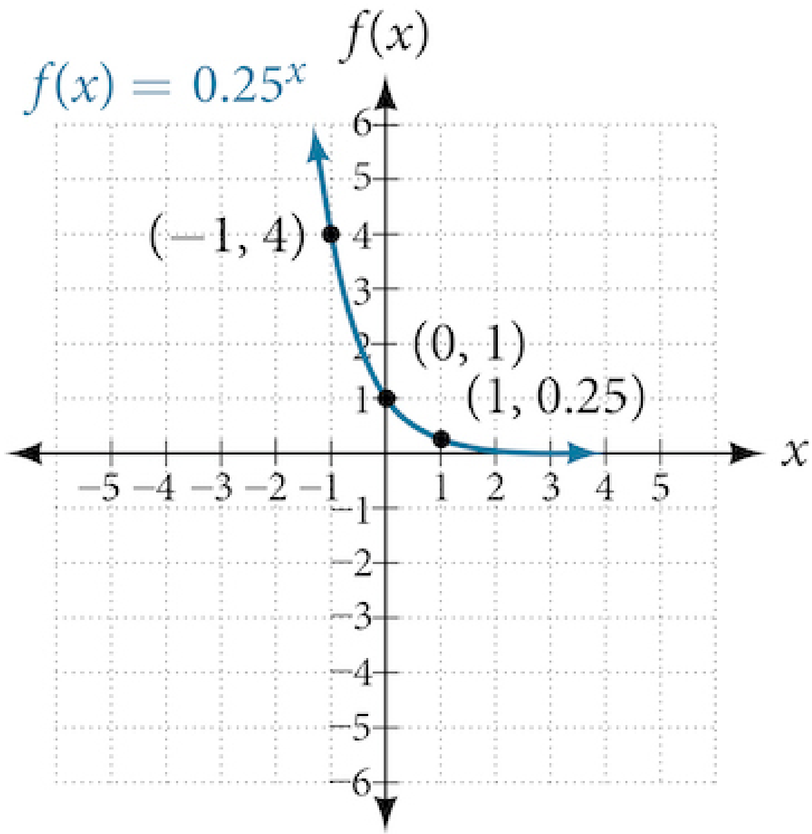



Figure 6.25: Exponential $(0.25)^x$ graph.

The domain is $\mathbb{R}$; the range is $(0,\infty)$; the horizontal asymptote is $y = 0$. ■

Notice again that a key factor for verifying exponential behavior will be confirming that for every change of 1 in the input, the value of the output conforms to the rule:

$$\frac{\text{New Value} - \text{Old Value}}{\text{Old Value}} = \text{Constant}$$

Notice in the example the constant value is $-0.75$ and $b = 0.25 = 1 + (-0.75)$.

## 6.3.5 Exponential Models

We have already explored some basic notions of exponential curves and representation. Now, we explore the real-world applications that make the more known models of exponential behavior, growth, and decay and their application to chemistry and economics.

### Modeling Exponential Growth and Decay

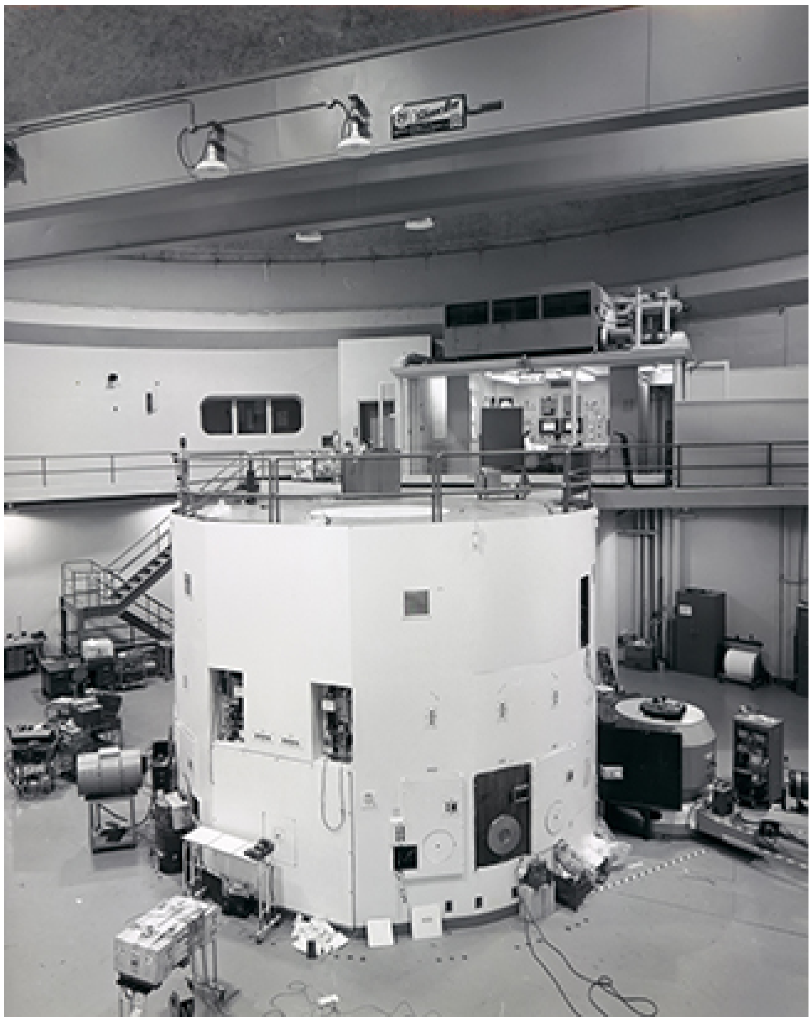


Figure 6.26: A nuclear research reactor inside the Neely Nuclear Research Center on the Georgia Institute of Technology campus (credit: Georgia Tech Research Institute).

In real-world applications, we need to model the behavior of a function. In mathematical modeling, we choose a familiar general function with properties that suggest that it will model the real-world phenomenon we wish to analyze. In the case of rapid growth, we may choose the exponential growth function. First, let's re-contextualize the exponential function for modeling purposes.

**Definition 6.3.5** An exponential function grows (or decays) by the same relative amount per unit time. The fractional growth rate:

$$\frac{\text{New Value} - \text{Old Value}}{\text{Old Value}} = r$$

Will now be called $r$, and any quantity $C$ growing (or decaying) at an exponential rate $r$ will behave as the exponential function.

$$C = C_0(1+r)^t \tag{6.4}$$

where

- $C$ is the value growing (or decaying) over amount of time $t$.
- $C_0$ is the initial value of the quantity $C$ at initial time $t = 0$.
- $r$ is the relative rate of growth for $C$.
- $t$ is the units of time.

### Population Growth or Decay

■ **Example 6.32** A population of bacteria doubles every hour. If the culture started with ten bacteria, state the exponential function governing the model.

**Solution:**
Notice our $C_0 = 10$ is quickly established. Similarly, bacteria doubles every hour means there is a 100% relative change each hour or $r = 1$. Clearly:

$$\begin{aligned} C &= C_0(1+r)^t \\ &= 10(1+1)^t = 10(2)^t \end{aligned}$$

■

■ **Example 6.33** A culture of bacteria is attacked via an antibacterial substance, reducing the population a 10% every minute. If the culture started with 10000 bacteria, state the exponential function governing the model.

**Solution:**
Notice $C_0 = 10000$ is established as our initial population, Now since the bacteria is reducing by a 10% relative change each minute or $r = -0.1$. Clearly:

$$\begin{aligned} C &= C_0(1+r)^t \\ &= 10000(1-0.1)^t = 10000(0.9)^t \end{aligned}$$

■

### Applying the Compound-Interest Formula

We use compound interest in savings instruments in which earnings are continually reinvested, such as mutual funds and retirement accounts. The term compounding refers to interest earned not only on the original value but on the accumulated value of the account.

The annual percentage rate (APR) of an account, also called the nominal rate, is the yearly interest rate earned by an investment account. The term nominal is used when the compounding occurs several times other than once yearly. When interest is compounded more than once a year, the effective interest rate is more significant than the nominal rate; this is a powerful tool for investing.

We can calculate the compound interest using the compound interest formula, which is an exponential function of the variables time $t$, principal $P$, APR $r$, and number of compounding periods in a year $n$.

For example, observe Table 6.27, which shows the result of investing $\$1,000$ at $10\%$ for one year. Notice how the value of the account increases as the compounding frequency increases.

| Frequency | Value after 1 year |
|---|---|
| Annually | $1100 |
| Semiannually | $1102.50 |
| Quarterly | $1103.81 |
| Monthly | $1104.71 |
| Daily | $1105.16 |

Figure 6.27: Compound Interest at different times.

**Definition 6.3.6** Compound interest can be calculated using the formula

$$A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$$

where

- $A(t)$ is the account value.
- $t$ is measured time in years.
- $P$ is the starting amount of the account, often called the principal.
- $r$ is the annual percentage rate (APR) expressed as a decimal, and
- $n$ is the number of compounding periods in one year.

The compound interest formula can be considered an exponential model if we consider the base $b = \left(1 + \frac{r}{n}\right)^n$ in the function $A(t) = b^t$.

■ **Example 6.34 — Finding worth after a specific time.** If we invest $\$3,000$ in an investment account paying 3% interest compounded quarterly, how much will the account be worth in 10 years?

**Solution:**

Because we are starting with $\$3,000$, $P = 3000$. Our interest rate is 3%, so $r = 0.03$. Because we are compounding quarterly, we are compounding 4 times per year, so n=4. We want to know the value of the account in 10 years, so we are looking for $A(10)$, the value when $t = 10$.

$$\begin{aligned} t &= \frac{\ln(r)}{-0.000121} \quad \text{Use the general form of the equation} \\ &= \frac{\ln(0.20)}{-0.000121} \quad \text{We substitute for } r. \\ &\approx 13301 \quad \text{Round to the nearest year} \end{aligned}$$

The account will be worth about $\$4,045.05$ in 10 years. ■

■ **Example 6.35 — Using the Compound Interest Formula to Solve for the principal.** A 529 Plan is a college-savings plan that allows relatives to invest money to pay for a child's future college tuition; the account grows tax-free. Lily wants to set up a 529 account for her new granddaughter and wants the account to grow to $\$40,000$ over 18 years. She believes the account will earn 6% compounded semi-annually (twice a year). To the nearest dollar, how much will Lily need to invest in the account now?

**Solution:**

The nominal interest rate is 6%, so $r = 0.06$. Interest is compounded twice a year, so $n = 2$.

We want to find the initial investment, $P$, needed so that the account's value will be worth $\$40,000$ in 18 years. Substitute the given values into the compound interest formula, and solve for $P$.

$$\begin{aligned} A(t) &= P\left(1+\frac{r}{n}\right)^{nt} \quad \text{Use the compound interest formula} \\ 40000 &= P\left(1+\frac{0.06}{2}\right)^{2(18)} \quad \text{We substitute for A, r, n, and t} \\ 40000 &= P(1.03)^{36} \quad \text{Simplify} \\ \frac{40000}{(1.03)^{36}} &\approx 13,801 \approx P \quad \text{Isolate and approximate} \end{aligned}$$

Lily will need to invest $\$13,801$ to have $\$40,000$ in 18 years. ■

### The number $e$ and continuous growth

As we saw earlier, the amount earned on an account increases as the compounding frequency increases. Figure 6.28 shows that the rise from annual to semi-annual compounding is more significant than the increase from monthly to daily compounding; this might lead us to ask whether this pattern will continue.

Examine the value of \$1 invested at 100% interest for one year, compounded at various frequencies, listed in figure 6.28.

| Frequency | $A(n) = \left(1 + \frac{1}{n}\right)^n$ | Value |
|---|---|---|
| Annually | $\left(1 + \frac{1}{1}\right)^1$ | \$2 |
| Semiannually | $\left(1 + \frac{1}{2}\right)^2$ | \$2.25 |
| Quarterly | $\left(1 + \frac{1}{4}\right)^4$ | \$2.441406 |
| Monthly | $\left(1 + \frac{1}{12}\right)^{12}$ | \$2.613035 |
| Daily | $\left(1 + \frac{1}{365}\right)^{365}$ | \$2.714567 |
| Hourly | $\left(1 + \frac{1}{8760}\right)^{8760}$ | \$2.718127 |
| Once per minute | $\left(1 + \frac{1}{525600}\right)^{525600}$ | \$2.718279 |
| Once per second | $\left(1 + \frac{1}{31536000}\right)^{31536000}$ | \$2.718282 |

Figure 6.28: Compound Interest for a dollar

These values appear to be approaching a limit as $n$ increases without a bound. In fact, as $n$ gets larger and larger, the expression $\left(1 + \frac{1}{n}\right)^n$ approaches a number used so frequently in mathematics that it has its own name: the letter $e$. This value is an irrational number, meaning its decimal expansion goes on forever without repeating. Its approximation to six decimal places is shown below.

**Definition 6.3.7** The letter $e$ represents the irrational number

$$\left(1 + \frac{1}{n}\right)^n$$

The letter $e$ is a base for many real-world exponential models. The mathematician Leonhard Euler named the number, who first investigated many of its properties. To work with base e, we use the approximation, $e \approx 2.718282$.

So far, we have worked with rational bases for exponential functions. However, $e$ is used as the base for exponential functions for most real-world phenomena. Exponential models that use $e$ as the base are called continuous growth or decay models. We see these models in finance, computer science, and sciences, such as physics, toxicology, and fluid dynamics.

**Definition 6.3.8** For all real numbers $t$, and all positive numbers $a$ and $r$, continuous growth or decay is represented by the formula:

$$A(t) = ae^{rt}$$

where

- $a$ is the initial value.
- $r$ is the growth or interest rate per unit of time.
- and $t$ is the elapsed time.

If $r > 0$, then the formula represents continuous growth. If $r < 0$, then the formula represents ongoing decay.

For business applications, the continuous growth formula is called the continuous compounding formula and takes the form.

$$A(t) = Pe^{rt}$$

where

- $P$ is the principal of the initial invested
- $r$ is the growth or interest rate per unit of time.
- and $t$ is the period or term of investment.

**Example 6.36 — Using the Compound Interest Formula to Solve for the principal.** A person invested \$1,000 in an account, earning a nominal 10% yearly compounded continuously. How much was in the account at the end of one year?

**Solution:**

Since the account is growing in value, this is a continuous compounding problem with a growth rate of $r = 0.10$. The initial investment was \$1,000, so $P = 1000$. We use the continuous compounding formula to find the value after $t = 1$ year:

$$\begin{aligned} A(t) &= Pe^{rt} \quad \text{Use the continous compound interest formula} \\ &= 1000e^{0.1} \quad \text{We substitute for P, r, and t} \\ &\approx 1105.17 \quad \text{Use calculator and approximate} \end{aligned}$$

The account is worth \$1,105.17 after one year. ■

### Half-life and Doubling Time

We now turn to exponential decay. One of the standard terms associated with exponential decay is the half-life, the length of time it takes an exponentially decaying quantity to decrease to half its original amount. Every radioactive isotope has a half-life, and the process describing the exponential decay of an isotope is called radioactive decay.

To find the half-life of a function describing exponential decay, solve the following equation:

$$\frac{1}{2}A_0 = A_0e^{kt}$$

Then the half-life depends only on the constant $k$ and not on the starting quantity $A_0$. The formula is derived as follows:

$$\begin{aligned}
\frac{1}{2}A_0 &= A_0e^{kt} \\
\frac{1}{2} &= e^{kt} \text{ Divide by } A_0 \\
\ln\left(\frac{1}{2}\right) &= kt \text{ Take natural log} \\
-\frac{\ln(2)}{k} &= t \text{ Simplify and divide by } k
\end{aligned}$$

■ **Example 6.37** The half-life of carbon-14 is 5,730 years. Express the amount of carbon-14 remaining as a function of time, t.

**Solution:**
This formula is derived as follows:

$$\begin{aligned}
A &= A_0e^{kt} \text{ The continuous growth formula} \\
0.5A_0 &= A_0e^{5730k} \text{ We substitute using the half life time and } A_0 \\
\ln(0.5) &= 5730k \text{ Divide by } A_0 \text{ and take } \ln \\
\frac{\ln(0.5)}{5730} &= k \text{ we find the value for } k
\end{aligned}$$

The function that describes this continuous decay is

$$f(t) = A_0e^{\frac{\ln(0.5)}{5730t}}.$$

We observe that the coefficient of $t$, $\frac{\ln(0.5)}{5730t} \approx -1.2097 \times 10^{-4}$ is negative, as expected in the case of exponential decay. ■

In general exponential growth leads to repeated doubling, and exponential decay will produce repeated halving. Usually, when modeling, we have the relative growth rate, but sometimes we may instead have the time required for doubling or halving.

■ **Example 6.38** Consider a population that has 5000 members and grows with a doubling time of 5 years.

- In five years, after one doubling time the population increases by a factor of two, or $2(5000) = 10000$.
- In ten years, or after two doubling times the populations increases by a factor of $2^2$ or $4(5000) = 20000$.
- In fifteen years we have three doubling and an increase factor of $2^3$ or $8(5000) = 40000$.
- If we consider $T_D$ as the time for doubling, the number of doublings at time $t$ is $\frac{t}{T_D}$ and the population after time $t$ would be the initial population multiplied by $2^{\frac{t}{T_D}}$

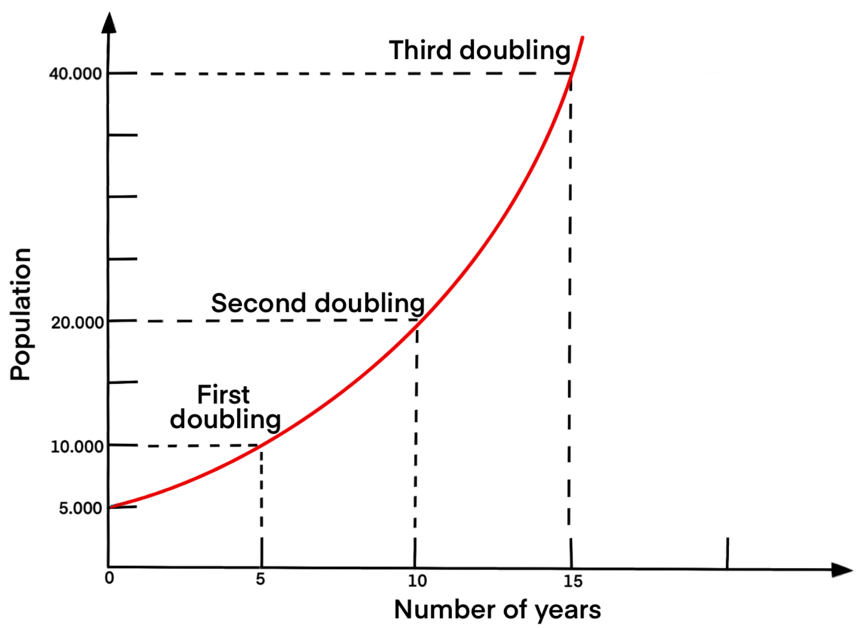



Figure 6.29: Doubling time in a exponential

■

Now with this idea, we can define the doubling and halving time more clearly.

**Definition 6.3.9 — Halving and Doubling Time.** The time required for each halving in exponential decay is called **Halving time**.

After a time $t$, an exponentially decaying quantity with a halving time of $T_H$ decreases in size by a **factor** of

$$\left(\frac{1}{2}\right)^{(t/T_H)}.$$

The new value of the growing quantity is related to its initial value (at $t = 0$) by

$$\text{New value} = \text{Initial value} \times \left(\frac{1}{2}\right)^{(t/T_H)}$$

The time required for each doubling in exponential growth is called **Doubling time**.

After a time $t$, an exponentially growing quantity with a doubling time of $T_D$ increases in size by a **factor** of

$$2^{(t/T_D)}$$

.

The new value of the growing quantity is related to its initial value (at $t = 0$) by

$$\text{New value} = \text{Initial value} \times 2^{(t/T_D)}$$

■ **Example 6.39 — Protein and Glucose in Bacterial Meningitis** .
An experimental substance creates a exponential growth in CSF protein level in a Bacterial Meningitis culture. At the same time, a exponential decay is observed at the CSF glucose levels.

If the protein level starts at 20 mg/dL and the glucose levels at 70 mg/dL calculate the new values if $T_D = 1.5$ hours and $T_H = 2.0$ hours, and 5 hours have passed.

**Solution:**

The new value of the growing protein after 5 hours.

$$\text{New value} = 20 \times (2)^{(5/1.5)} \approx 201.587 \ \text{ mg/dL}$$

The new value of the decaying glucose after 5 hours.

$$\text{New value} = 70 \times \left(\frac{1}{2}\right)^{(5/1.5)} \approx 6.9448 \ \text{ mg/dL}$$

Notice at this point is easy to obtain the growing and decay rates using the relative change formula.

■

A natural question that arises is if we know at what percentage rate a quantity is growing or decaying, how can we know the doubling time or the half-life, respectively?

First, we can try a heuristic method:

**Proposition 6.3.1 — Rule of 70**

**Exponential Growth or Decay**: Suppose a quantity is growing or decaying exponentially at $P\%$ per time period.

- **Approximate Half-Life Formula (The Rule of 70)**: The half-life or the doubling time is approximately
$$T_H \approx \frac{70}{P} \approx T_D$$
This approximation works best for low decay rates and breaks down for growth or decay rates over about 15%.

The rule of 70 is used to make quick approximations about doubling or halving times and is very popular in economic settings.

Now, if we want a more precise approximation, we can follow the half-life derivation idea at the start of this subsection, and we arrive at two similar results for halving and doubling time.

**Proposition 6.3.2 — Exact Halving and Doubling Times**

**Exponential Decay**: Suppose a quantity is decaying exponentially at a rate of $P\%$ per time period.

- **Exact Half-Life Formula**: For more precise work, use the exact formula. This uses the fractional decay rate, $r = P/100$. For an exponentially decaying quantity, the half-life is
$$T_H = -\frac{\ln 2}{\ln(1-r)}$$

**Exponential Growth**: Suppose a quantity is growing exponentially at a rate of $P\%$ per time period.

- **Exact Double Time Formula**: For more precise work, use the exact formula. This uses the fractional growth rate, $r = P/100$. For an exponentially growing quantity, the doubling time is
$$T_D = \frac{\ln 2}{\ln(1+r)}$$

### Radiocarbon Dating

The formula for radioactive decay is essential in radiocarbon dating, which calculates the approximate date a plant or animal died. Radiocarbon dating was discovered in 1949 by Willard Libby, who won a Nobel Prize for his discovery. It compares the difference between the ratio of two isotopes of carbon in an organic artifact or fossil to the ratio of those two isotopes in the air. It is believed to be accurate to about a 1% error for plants or animals that died within the last $60,000$ years.

Carbon-14 is a radioactive isotope of carbon that has a half-life of 5,730 years. It occurs in small quantities in the carbon dioxide in the air we breathe. Most of the carbon on Earth is carbon-12, which has an atomic weight of 12 and is not radioactive. Scientists have determined the ratio of carbon-14 to carbon-12 in the air for the last $60,000$ years, using tree rings and other organic samples of available dates—although the ratio has changed slightly over the centuries.

As long as a plant or animal is alive, the ratio of the two isotopes of carbon in its body is close to the ratio in the atmosphere. When it dies, the carbon-14 in its body decays and is not replaced. We can approximate the date the plant or animal died by comparing the ratio of carbon-14 to carbon-12 in a decaying sample to the known ratio in the atmosphere. Since the half-life of carbon-14 is $5,730$ years, the formula for the amount of carbon-14 remaining after t years is

$$A = A_0 e^{\frac{\ln(0.5)}{5730t}}.$$

where:

- $A$ is the amount of carbon-14 remaining.
- $A_0$ is the amount of carbon-14 when the plant or animal began decaying.
- To find the age of an object, we solve this equation for $t$:

$$t = \frac{\ln(\frac{A}{A_0})}{-0.000121}$$

Out of necessity, we neglect here the many details that a scientist considers when doing carbon-14 dating, and we only look at the basic formula. The ratio of carbon-14 to carbon-12 in the atmosphere is approximately $0.0000000001\%$. Let $r$ be the ratio of carbon-14 to carbon-12 in the organic artifact or fossil to be dated, determined by a liquid scintillation method. From the equation $A \approx A_0 e^{-0.000121t}$ we know the ratio of the percentage of carbon-14 in the object we are dating to the percentage of carbon-14 in the atmosphere is $r = \frac{A}{A_0} \approx e^{-0.000121t}$. We solve this equation for $t$ to get

$$t = \frac{\ln(r)}{-0.000121}$$

■ **Example 6.40** A bone fragment is found that contains 20% of its original carbon-14. To the nearest year, how old is the bone?.

**Solution:**
We substitute $20\% = 0.20$ for $k$ in the equation and solve for $t$:

$$\begin{aligned} t &= \frac{\ln(r)}{-0.000121} \quad \text{Use the general form of the equation} \\ &= \frac{\ln(0.20)}{-0.000121} \quad \text{We substitute for } r. \\ &\approx 13301 \quad \text{Round to the nearest year} \end{aligned}$$

The bone fragment is about 13,301 years old.

The instruments that measure the percentage of carbon-14 are susceptible, and, as we mentioned above, a scientist will need to do much more work than we did to be satisfied. Even so, carbon dating is only accurate to about 1%, so this age should be given as $13,301$ years $\pm 1\%$ or $13,301$ years $\pm 133$ years. ■

### Newton's Law of Cooling

Exponential decay can also be applied to temperature. When a hot object is left in the surrounding air at a lower temperature, the object's temperature will decrease exponentially, leveling off as it approaches the surrounding air temperature. On a graph of the temperature function, the leveling off will correspond to a horizontal asymptote at the surrounding air temperature. Unless the room temperature is zero, this will correspond to a vertical shift of the generic exponential decay function. This translation leads to Newton's Law of Cooling, the scientific formula for temperature as a function of time as an object's temperature is equalized with the ambient temperature.

$$T(t) = ae^{kt} + T_s$$

where:

- $t$ is time.
- $A$ is the difference between the initial temperature of the object and the surroundings.
- $k$ is a constant, the continuous rate of cooling of the object.
- $T_s$ is usually the ambient temperature.

To find the time or temperature, use Newton's Law of Cooling.

■ **Example 6.41** A cheesecake is taken out of the oven with an ideal internal temperature of 165°F and is placed into a 35°F refrigerator. After 10 minutes, the cheesecake cooled to 150°F. If we must wait until the cheesecake has cooled to 70°F before we eat it, how long will we have to wait?

**Solution:**

Because the surrounding air temperature in the refrigerator is 35 degrees, the cheesecake's temperature will decay exponentially toward 35, following the equation.

$$T(t) = Ae^{kt} + T_s$$

We know the initial temperature was 165, so $T(0) = 165$.

$$\begin{aligned} 165 &= Ae^{k(0)} + 35 \;\; \text{Substitute}(0,165) \\ A &= 130 \;\; \text{Solve for } A \end{aligned}$$

We were given another data point, $T(10) = 150$, which we can use to solve for $k$.

$$\begin{aligned} 150 &= 130e^{k(10)} + 35 \;\; \text{Substitute}(10,150) \\ 115 &= 130e^{k(10)} \;\; \text{Substract 35.} \\ \frac{115}{130} &= e^{k(10)} \;\; \text{Divide 130.} \\ \ln\left(\frac{115}{130}\right) &= k(10) \;\; \text{Apply ln.} \\ k &= \ln\left(\frac{115}{130}\right) \approx -0.0123 \;\; \text{Divide by coefficient of } k. \end{aligned}$$

This gives us the equation for the cooling of the cheesecake:

$$T(t) = 130e^{-0.0123t} + 35$$

Now we can solve the time it will take for the temperature to cool to 70 degrees.

$$\begin{aligned} 70 &= 130e^{-0.0123t} + 35 \;\; \text{Substitute } T(t) \text{ in } 70 \\ 35 &= 130e^{-0.0123t} \;\; \text{Substract 35.} \\ \frac{35}{130} &= e^{-0.0123t} \;\; \text{Divide 130.} \\ \ln\left(\frac{35}{130}\right) &= -0.0123t \;\; \text{Apply ln.} \\ t &= \ln\left(\frac{35}{130}\right) / -0.0123 \approx 106.68 \;\; \text{Divide by coefficient of } t. \end{aligned}$$

It will take about 107 minutes, or one hour and 47 minutes, for the cheesecake to cool to 70°F.

■

### 6.3.6 Exercises 6.3

**Problem 6.46** Explain why the values of an increasing exponential function will eventually overtake the importance of an increasing linear function.

**Problem 6.47**

Given a formula for an exponential function, is it possible to determine whether the function grows or decays exponentially just by looking at the formula? Explain.

**Problem 6.48** For the following exercises, identify whether the statement represents an exponential function. Explain.

- The average annual population increase of a pack of wolves is 25.
- A population of bacteria decreases by a factor of $\frac{1}{8}$ every 24 hours.
- The value of a coin collection has increased by 3.25% annually over the last 20 years.
- For each training session, a personal trainer charges his clients \$5 less than the previous training session.
- The height of a projectile at time $t$ is represented by the function
$$h(t) = -4.9t^2 + 18t + 40.$$

**Problem 6.49** Each year $t$, the population of a forest of trees is represented by the function.

$$A(t) = 115\,(1.025)^t$$

In a neighboring forest, the population of the same type of tree is represented by the function

$$B(t) = 82\,(1.029)^t$$

- Which forest's population is growing at a faster rate?
- Which forest had a greater number of trees initially? By how many?
- Assuming the population growth models continue to represent the growth of the forests, which forest will have a greater number of trees after 20 years? By how many? Then what about after 100 years? By how many?

**Problem 6.50** The following exercises determine whether the equation represents exponential growth, exponential decay, or neither. Explain.

- $y = 300(1-t)^5$
- $y = 220(1.06)^x$
- $y = 16.5(1.025)^{\frac{1}{x}}$
- $y = 11,701(0.97)^t$

**Problem 6.51** For the following exercises, find the formula for an exponential function that passes through the two points.

- $(0,6)$ and $(3,750)$
- $(0,2000)$ and $(2,20)$
- $(-1,\frac{3}{2})$ and $(3,24)$
- $(-2,6)$ and $(3,1)$
- $(3,1)$ and $(5,4)$

**Problem 6.52** The fox population in a particular region has an annual growth rate of 9%per year. In the year 2012, there was 23,900 fox counted in the area. What is the fox population predicted to be in the year 2020?

**Problem 6.53** A scientist begins with 100 milligrams of a radioactive substance that decays exponentially. After 35 hours, 50mg of the substance remains. How many milligrams will remain after 54 hours?

**Problem 6.54** The wolf population in Alaska has an annual growth rate of 7%per year. In the year 2010, there were 3,900 wolves counted in the area. What is the wolf population predicted to be in the year 2025?

**Problem 6.55** The table shows the number of monthly active users of Facebook (the number of people who use Facebook at least once a month) over four years, measured in March each year.

| March-year | 2009 | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|---|
| Active Users (Millions) | 197 | 431 | 671 | 902 | 1137 |
| Absolute Difference | | | | | |
| Relative Difference | | | | | |

- Fill in the third row of the table showing the absolute change in the number of active monthly users.
- Fill in the fourth row of the table showing the percent change in the number of active monthly users.
- Is the increase in Facebook users linear or exponential? Justify your answer.
- If we suppose the trend in the monthly active users of Facebook continues as shown in the table, what is the number of Facebook users in March 2014?

**Problem 6.56** For the following exercises, determine whether the table could represent a function that is linear, exponential or neither.

- The function given by:

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $f(x)$ | 70 | 40 | 10 | -20 |

- The function given by:

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $h(x)$ | 70 | 49 | 34.3 | 24.01 |

Find a function that passes through the points if it appears to be exponential.

**Problem 6.57** A researcher is studying the growth of cancerous cells in a tumor. She records the date in the following table:

| Week | Number of cancerous cells |
|---|---|
| 0 | 2000 |
| 5 | 2492 |
| 10 | 3106 |
| 15 | 3871 |
| 20 | |
| 25 | |

1. Is the number of cancerous cells growing linearly? exponentially? Support your conclusion with any relevant calculations.

2. Complete the missing values in the table.

**Problem 6.58** After a certain number of years, the value of an investment account is represented by the equation.

$$A = 10250\left(1+\frac{0.04}{12}\right)^{120}$$

- What is the value of the account?

- What was the initial deposit made to the account?

- How many years had the account accumulating interest?

**Problem 6.59** An account is opened with an initial deposit of $\$6,500$ and earns $3.6\%$ interest compounded semi-annually. What will the account be worth in 20 years?

- How much more would the account be worth if the interest were compounding weekly?

- How much more would be worth if you increase the principal to $\$7,500$

**Problem 6.60** Suppose an investment account is opened with an initial deposit of $\$12,000$ earning $7.2\%$ interest compounded continuously. How much will the account be worth after 30 years?

**Problem 6.61** For the following exercises, determine whether the table could represent a function that is linear, exponential or neither. Find a function that passes through the points if it appears to be exponential.

- The function given by:

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $m(x)$ | 80 | 61 | 42.9 | 25.61 |

- The function given by:

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $f(x)$ | 10 | 20 | 40 | 80 |

- The function given by:

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $g(x)$ | -3.25 | 2 | 7.25 | 12.5 |

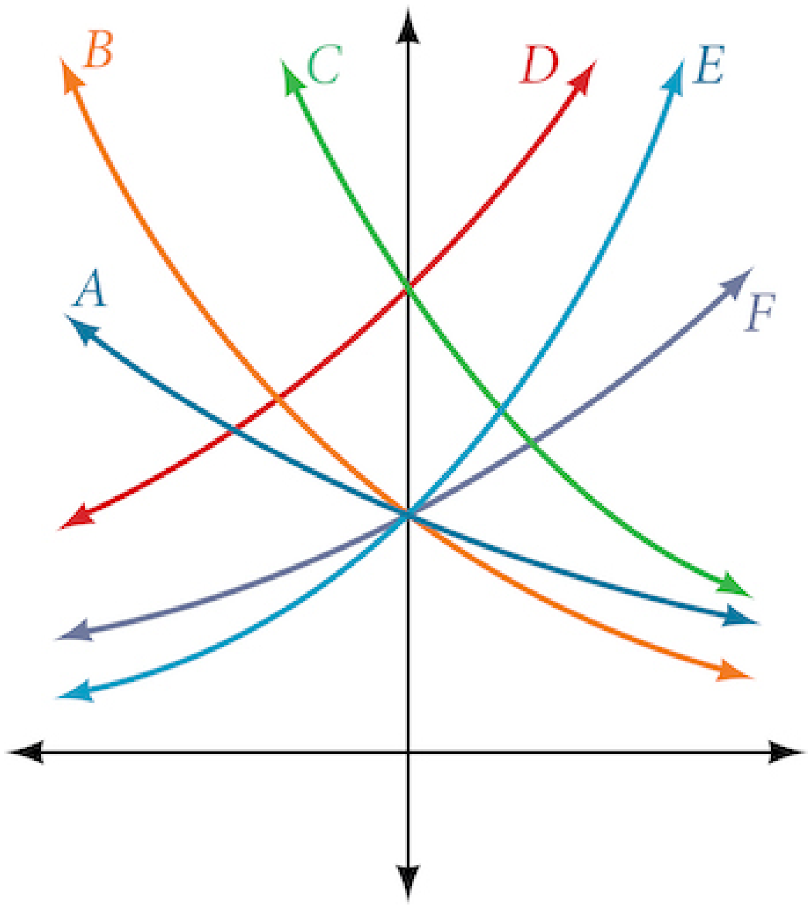
**Problem 6.62** The following exercises determine whether the equation represents exponential growth, exponential decay, or neither. Explain.

- $y = 3742e^{0.75t}$
- $y = (2.25)^{-2t}$
- $y = 150e^{3.25/t}$
- $y = (t)^2$

**Problem 6.63** For the following exercises, evaluate each function. Round answers to four decimal places, if necessary.

- $f(x) = 2(5)^x$ for $f(-3)$
- $f(x) = -(4)^{2x+3}$ for $f(-1)$
- $f(x) = e^x$ for $f(3)$
- $f(x) = 2.7(4)^{-x+1}$ for $f(-2)$

**Problem 6.64** For the following exercises, match each function with one of the graphs in




- $f(x) = 2(0.69)^x$
- $f(x) = 2(1.28)^x$
- $f(x) = 2(0.81)^x$
- $f(x) = 4(1.28)^x$
- $f(x) = 2(1.59)^x$
- $f(x) = 4(0.69)^x$

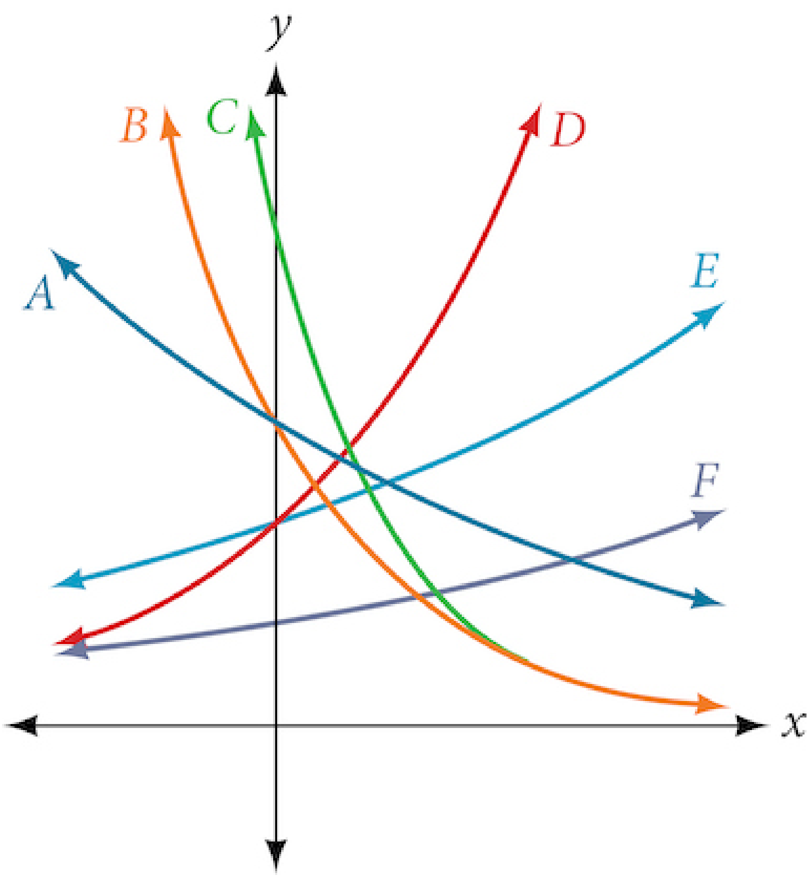
**Problem 6.65** A car was valued at \$38,000 in the year 2007. By 2013, the value had depreciated to \$11,000. If the car's value continues to drop by the same percentage, what will it be worth by 2017?

**Problem 6.66** Jamal wants to save \$54,000 for a down payment on a home. How much will he need to invest in an account with 8.2% APR, compounding daily, to reach his goal in 5 years?

**Problem 6.67** Kyoko has $10,000 that she wants to invest. Her bank has several investment accounts to choose from, all compounding daily. Her goal is to have $15,000 by the time she finishes graduate school in 6 years.

To the nearest hundredth of a percent, what should her minimum annual interest rate be to reach her goal? (Hint: solve the compound interest formula for the interest rate.)

**Problem 6.68** For the following exercises, use the graphs shown below. All have the form $f(x) = ab^x$.




- Which graph has the largest value for b?
- Which graph has the smallest value for b?
- Which graph has the largest value for a?
- Which graph has the smallest value for a?

**Problem 6.69** We open an investment account with an annual interest rate of 7% with an initial deposit of \$4,000. Compare the account's values after nine years when the interest is compounded annually, quarterly, monthly, and continuously.

**Problem 6.70** With what kind of exponential model would half-life be associated? What role does half-life play in these models?

**Problem 6.71** Determine whether the data from the table could best be represented as a function that is linear, exponential, or none. If possible, write a formula for a model that describes the data.

| $x$ | $f(x)$ |
|---|---|
| –2 | 0.694 |
| –1 | 0.833 |
| 0 | 1 |
| 1 | 1.2 |
| 2 | 1.44 |
| 3 | 1.728 |
| 4 | 2.074 |
| 5 | 2.488 |

**Problem 6.72** What is carbon dating? Why does it work? Give an example in which carbon dating would be useful.

**Problem 6.73**

With what kind of exponential model would doubling time be associated? What role does doubling time play in these models?

**Problem 6.74** Define Newton's Law of Cooling. Then name at least three real-world situations where we would apply Newton's Law of Cooling.

**Problem 6.75** The temperature of an object in degrees Fahrenheit after t minutes is represented by the equation

$$T(t) = 68e^{-0.0174t} + 72.$$

To the nearest degree, what is the object's temperature after one and a half hours?

**Problem 6.76** A scientist begins with 250 grams of a radioactive substance. After 250 minutes, the sample decayed to 32 grams. Rounding to five decimal places, write an exponential equation representing this situation. To the nearest minute, what is the half-life of this substance?

**Problem 6.77** For the following exercises, use this scenario: A biologist recorded a count of 360 bacteria present in culture after 5 minutes and 1000 bacteria present after 20 minutes.

- To the nearest day, how long will it take half of the Iodine-125 to decay?
- To the nearest whole number, what was the initial population in the culture?
- Rounding to six decimal places, write an exponential equation representing this situation. How long did it take the population to double to the nearest minute?

**Problem 6.78** Sketch the graph of $f(x) = 3^x$.

**Problem 6.79**

For the following exercises, use this scenario: A doctor prescribes 125 milligrams of a therapeutic drug that decays by about 30% each hour.

- To the nearest hour, what is the drug's half-life?
- Write an exponential model representing the amount of the drug remaining in the patient's system after $t$ hours. Then use the formula to find the amount of the drug that would remain in the patient's system after 3 hours. Round to the nearest milligram.
- Using the model found in the previous exercise, find $f(10)$ and interpret the result. Round to the nearest hundredth.

**Problem 6.80** Sketch the graph of $f(x) = \left(\frac{1}{4}\right)^x$.

**Problem 6.81** For the following exercises, use this scenario: A tumor is injected with 0.5 grams of Iodine-125, which has a decay rate of 1.15% per day.

- To the nearest day, how long will it take half of the Iodine-125 to decay?
- Write an exponential model representing the amount of the drug remaining in the patient's system after $t$ hours. Then use the formula to find the amount of the drug that would remain in the patient's system after 3 hours. Round to the nearest milligram.
- Write an exponential model representing the amount of Iodine-125 remaining in the tumor after t days. Then use the formula to find the amount of Iodine-125 that would remain in the tumor after 60 days. Round to the nearest tenth of a gram.

**Problem 6.82** A scientist begins with 250 grams of a radioactive substance. After 250 minutes, the sample decayed to 32 grams. Rounding to five decimal places, write an exponential equation representing this situation. To the nearest minute, what is the half-life of this substance?

**Problem 6.83** For the following exercises, use this scenario: A pot of warm soup with an internal temperature of 100° Fahrenheit was taken off the stove to cool in a 69° F room. After fifteen minutes, the internal temperature of the soup was 95° F.

- Use Newton's Law of Cooling to write a formula that models this situation.
- To the nearest minute, how long will it take the soup to cool to 80° F?
- To the nearest degree, what will the temperature be after two and a half hours?

**Problem 6.84** Suppose there is a population of fruit flies with a doubling time of 8 hours and an initial population of 7 fruit flies.

- Fill in the blanks:

  The population of fruit flies is ____________ (growing/decaying)____________ (linearly/exponentially).

- By what factor does the population grow in 24 hour? in 3 days?
- What is the population of fruit flies after 24 hours? after 3 days?

**Problem 6.85** The population of a threatened animal species is 1 million, but it is declining with a half-life of 20 years.

- Fill in the blanks:

  The population is ____________ (growing/decaying)____________ (linearly/exponentially).

- What fraction of the original population remains after 40 years? after 60 years?
- How many animals will be left after 30 years? after 40 years?

**Problem 6.86** A family of 10 termites invades your house, and its population increases at 20% per week.

1. Fill in the blanks:

   The population of termites is ____________ (growing/decaying)____________ (linearly/exponentially).

2. Compare the doubling times found with the approximate and the exact doubling time formulas.

3. How many termites will be in your house after eight weeks? After six months?

4. How many weeks does it take to have 5 thousand termites in your house? What about 10 thousand termites?

**Problem 6.87**

A particular antibiotic is administered to an individual. The table below shows information about the amount of drug in the individual's bloodstream, $A$, measured in mg, $t$ hours after taking the drug.

| t (hours) | 0 | 2 | 4 | 6 | 8 | 10 | 12 |
|---|---|---|---|---|---|---|---|
| A (mg) | 300.00 | 238.11 | 188.99 | 150.00 | 119.09 | 94.49 | 75.00 |

1. What was the initial amount of drug administered to this individual?

2. What is the half-life of this drug?

3. How long will it take for the amount of drug in the bloodstream to be 37.50 mg?

4. By approximately what percent does the amount of drug decrease every hour?

5. A second dose needs to be administered when the amount of drug in the bloodstream from the first dose equals 5 mg. How long does it take to administer the second dose?

**Problem 6.88** As of August 31, 2014, there were approximately 3685 cases of Ebola virus disease in the West African nations of Guinea, Liberia, and Sierra Leone. By October 6, the cumulative number of patients had grown to 7412.

| Date | Week | Total Cases |
|---|---|---|
| 1-Sep 2014 | 0 | 3685 |
| 8-Sep-2014 | 1 | 4238 |
| 15-Sep-2014 | 2 | 4873 |
| 22-Sep-2014 | 3 | 5604 |
| 29-Sep-2014 | 4 | 6445 |
| 6-Oct-2014 | 5 | |
| 13-Oct-2014 | 6 | |
| 20-Oct-2014 | 7 | |

- Explain why a linear function would not be a good model for this data set.
- By what percent does the number of cases increase between week 0 and week 1? Between week three and week 4?
- Complete the next two lines of the table and write a sentence describing how the number of cases is growing.
- Write a formula for $C(t)$, the total number of cases, where $t$ represents the number of weeks after September 1, 2014, and predict the number of cases on January 5, 2015 (week 18).

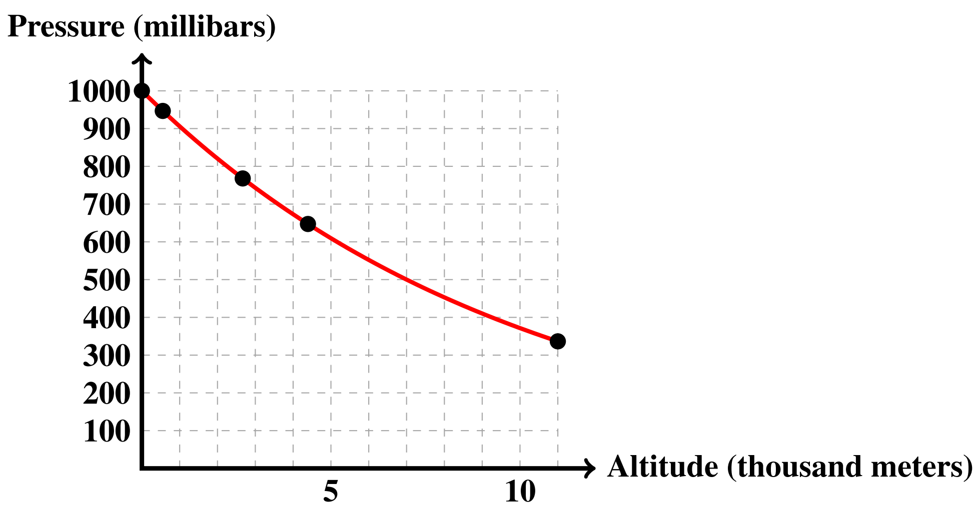
**Problem 6.89** The current population of the thick-billed parrot species is $\frac{1}{2}$ million, but it is declining with a half-life of 10 years.

- Fill in the blanks:

  The population is ____________ (growing/decaying)____________ (linearly/exponentially).
- What fraction of the original population remains after 20 years? after 40 years?
- How many animals will be left after 20 years? after 40 years?

**Problem 6.90** Let's explore the behavior of the function that relates altitude, $h$, in meters, and atmospheric pressure, $P(h)$, in millibars. Below is a table and a graph for a set of altitude values and pressure values.

| $h$ (meters) | $P(h)$ (millibars) | Description |
|---|---|---|
| 0 | 1000 | Sea level |
| 553 | 946.71 | CN Tower, Toronto, Ontario, Canada |
| 2,667 | 767.91 | Guadalupe Peak, highest point in Texas |
| 4,392 | 647.33 | Mt. Rainier, highest point in Washington state |
| 11,000 | 336.48 | Commercial airliner altitude |




1. Explain why a linear function would not be a good model for this set of data.
2. At which altitude is the air pressure equal to 500 millibars?
3. Write a formula for $P(h)$, the atmospheric pressure, in terms of $h$, the altitude.
4. What is the atmospheric pressure at the top of Mount Everest(8848 meters)? What is the atmospheric pressure at Death Valley National Park (86 meters below sea level)? Do your answers make sense?

## 6.4 Project In Modeling

When we learn new models, we like to know why they matter. In this activity, we will analyze data from a population of the United States and aim to determine what model fits better. We want to:

- Determine what function best models the growth of a population–linear, exponential, quadratic, or logarithmic.
- Use their models to predict what the U.S. population will look like in the future.

Using data from the U.S. Census Bureau, you will analyze U.S. population trends from 1990 through 2010. The data has been broken down by demographic, so you can see how the population trends of different groups compare.

| Real Year | Year | American Indian | Asian and Pacific Islander | Black | Hispanic | White (non-Hispanic) | Total Population |
|---|---|---|---|---|---|---|---|
| **1900** | 0 | 237,196 | 114,189 | 8,833,994 | 503,189 | 66,374,317 | 75,994,575 |
| **1910** | 10 | 265,683 | 146,863 | 9,827,763 | 797,994 | 81,043,248 | 91,972,266 |
| **1920** | 20 | 244,437 | 182,137 | 10,463,131 | 1,286,154 | 93,604,612 | 105,710,620 |
| **1930** | 30 | 332,397 | 264,766 | 11,891,143 | | | 122,775,046 |
| **1940** | 40 | 333,969 | 254,918 | 12,865,518 | 2,021,820 | 116,261,189 | 131,669,275 |
| **1950** | 50 | 343,410 | 321,033 | 15,042,286 | 3,231,409 | 131,805,405 | 150,697,361 |
| **1960** | 60 | 551,669 | 980,337 | 18,871,831 | 5,814,784 | 153,217,498 | 179,323,175 |
| **1970** | 70 | 795,110 | 1,526,401 | 22,539,362 | 8,920,940 | 169,622,593 | 203,210,158 |
| **1980** | 80 | 1,420,400 | 3,500,439 | 26,495,025 | 14,608,673 | 180,256,366 | 226,545,805 |
| **1990** | 90 | 1,959,234 | 7,273,662 | 29,986,060 | 22,354,059 | 188,128,296 | 248,709,873 |
| **2000** | 100 | 2,475,956 | 10,641,833 | 34,658,190 | 35,305,818 | 194,552,774 | 281,421,906 |
| **2010** | 110 | 2,932,248 | 15,214,265 | 38,929,319 | 50,477,594 | 196,817,552 | 308,745,538 |

### 6.4.1 Part I: The Statistic Approach.

Choose one demographic and analyze its population growth since 1990. Before you make your scatter plot, answer the following questions.

(a).Which demographic did you choose?

(b). What correlation do you see between time and population? (Positive or negative? Strong or weak?)

(c) Investigate the coefficient of determination R.

(d). Make your scatter plot, add a trendline, and display the equation and r-squared value of your trendline.

(e). What is the coefficient of determination (R-squared value) for the linear trendline?

(f). What is the R-squared value for the exponential trendline?

(g). What model fits your data best—linear or exponential? How do you know?

(h). Why do you think that model worked best?

(i). What is the equation for your trendline?

(j). Using the equation, predict how big the population will be in 2020 (120 years after 1900). Show your work.

(k). Using the equation, predict how big the population will be in 2050. Show your work.

(l). According to your function, when will the population reach 75,000,000?

### 6.4.2 Part I: The Modeling Approach.

Compare the population trends of two demographics since 1900. Instead of making just one trendline, you'll make two this time, one for each demographic.

(a). What is the coefficient of determination ($R$-squared value) for the "White (non-Hispanic)" linear trendline? For the exponential trendline? Circle the one that fits the data best—linear or exponential.

- Linear
- Exponential

(b). What is the equation for the trendline for the "White (non-Hispanic)" data?

(c). What is the $R$-squared value for the "Non-white" linear trendline? For the exponential trendline? Circle the one that fits the data best—linear or exponential.

- Linear
- Exponential

(d). What is the equation for the trendline for the "Non-white" data?

(e). The best model was different for the two demographics. Why do you think linear worked best for one and exponential worked best for the other?

(f). Using your trendline equations, complete the table below.

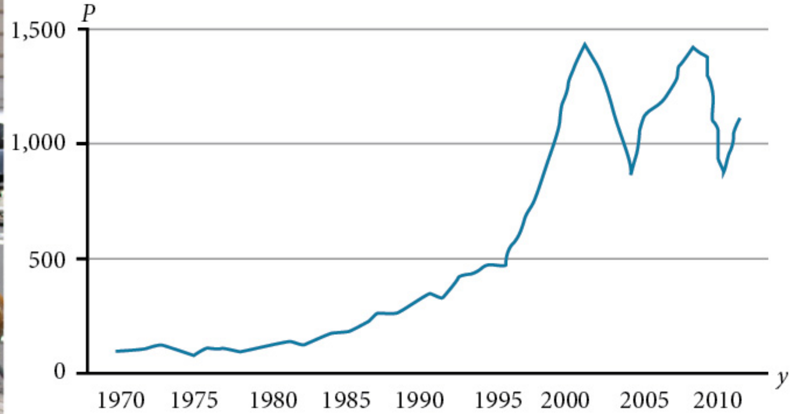
| Year | White (Non-Hispanic) | Non-White |
|---|---|---|
| 2020(120) | | |
| 2020(130) | | |
| 2020(140) | | |
| 2020(150) | | |
| 2020(160) | | |
| 2020(170) | | |

(f). Calculate relative difference and Absolute differences for both “White (non-hispanic)” and "Non-White data".

(g) If we were to approximate the values of both the relative differences and absolute differences for this data, will we obtain the notion of a Linear or Exponential Model?

(h) Your findings in the previous questions match with what the R-squares says for the fitting of the models?

(i) According to your calculations, when will non-whites outnumber whites (non-hispanic), if they do?




# Bibliography


[Abr15] Jay Abramson. *College Algebra*. 1st edition. Volume 3. 2. Open Stax, Jan. 2015, pages 159–463 (cited on page 313).

[Bar13] Susan Dean Barbara Illowsky. *Introductory Statistics*. Open Stax, 2013 (cited on page 225).

[BD20] Tifton. Bainbridge. Moultrie. Blakely and Donaldson. *Quantitative Reasoning Workbook*. 1st edition. Abraham Baldwin Agricultural College, 2020 (cited on page 191).

[Car86] Lewis Carroll. *Symbolic Logic*. 1.4. Mac Millan, 1986 (cited on page 145).

[Cle16] Matthew J. Van Cleave. *Introduction to Critical Thinking*. 1.4. Open Source, 2016 (cited on page 145).

[Dia17] Darlene Diaz. *Math for Liberal Arts Students*. 1.99. My Open Math, 2017 (cited on page 8).

[Lan21] David Lane. *Introductory Statistics*. Libre Texts, 2021. URL: `https://stats.libretexts.org/@go/page/2074` (cited on page 225).

[Lip17] David Lippman. *Math in Society*. 2.5. My Open Math, 2017 (cited on page 8).

[Mar20] Honeycut-Mathis Marecek Anthony-Smith. *Pre-Algebra*. 2nd edition. Open Stax, 2020 (cited on page 8).

[SZ21] Shafer and Zhang. *Introductory Statistics*. Libre Texts, 2021. URL: `https://stats.libretexts.org/@go/page/653` (cited on page 225).

[Wag09] David Arnold. Bruce Wagner. *Prealgebra Textbook*. 2nd edition. Lulu, 2009 (cited on page 8).