MATH FOR LIBERAL ARTS

Ashley McHale
Las Positas College

LibreTexts™

Math for Liberal Arts

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

This text was compiled on 06/09/2025

TABLE OF CONTENTS

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing****.*

CHAPTER OVERVIEW

1: Problem Solving

Thumbnail: pixabay.com/photos/tax-form-irs-tax-taxes-finance-4080693/

1.1: Introduction

In previous math courses, you've no doubt run into the infamous "word problems." Unfortunately, these problems rarely resemble the type of problems we actually encounter in everyday life. In math books, you usually are told exactly which formula or procedure to use, and are given exactly the information you need to answer the question. In real life, problem solving requires identifying an appropriate formula or procedure, and determining what information you will need (and won't need) to answer the question.

In this chapter, we will review several basic but powerful algebraic ideas: percents, rates, and proportions. We will then focus on the problem solving process, and explore how to use these ideas to solve problems where we don't have perfect information.

1.2: Percents

In the 2004 vice-presidential debates, Edwards's claimed that US forces have suffered "90% of the coalition casualties" in Iraq. Cheney disputed this, saying that in fact Iraqi security forces and coalition allies "have taken almost 50 percent" of the casualties[1]. Who is correct? How can we make sense of these numbers?

Percent literally means "per 100," or "parts per hundred." When we write 40%, this is equivalent to the fraction $\frac{40}{100}$ or the decimal 0.40. Notice that 80 out of 200 and 10 out of 25 are also 40%, since $\frac{80}{200}=\frac{10}{25}=\frac{40}{100}$.

Example 1

243 people out of 400 state that they like dogs. What percent is this?

Solution

$\frac{243}{400}=0.6075=\frac{60.75}{100}$. This is 60.75%.

Notice that the percent can be found from the equivalent decimal by moving the decimal point two places to the right.

Example 2

Write each as a percent: a) $\frac{1}{4}$ b) 0.02 c) 2.35

Solution

a) $\frac{1}{4}=0.25=25\%$ b) $0.02=2\%$ c) $2.35=235\%$

Percents

If we have a *part* that is some *percent* of a *whole*, then

$\text{percent} = \frac{\text{part}}{\text{whole}}$, or equivalently, $\text{part} = \text{percent} \cdot \text{whole}$

To do the calculations, we write the percent as a decimal.

Example 3

The sales tax in a town is 9.4%. How much tax will you pay on a $140 purchase?

Solution

Here, $140 is the whole, and we want to find 9.4% *of* $140. We start by writing the percent as a decimal by moving the decimal point two places to the left (which is equivalent to dividing by 100). We can then compute:

$\text{tax}=0.094(140)=\$13.16$in tax.

Example 4

In the news, you hear "tuition is expected to increase by 7% next year." If tuition this year was $1200 per quarter, what will it be next year?

Solution

The tuition next year will be the current tuition plus an additional 7%, so it will be 107% of this year's tuition:

$\$1200(1.07)=\$1284$

Alternatively, we could have first calculated 7% of $1200: $\$1200(0.07)=\$84$

Notice this is *not* the expected tuition for next year (we could only wish). Instead, this is the expected *increase*, so to calculate the expected tuition, we'll need to add this change to the previous year's tuition:

$\$1200+\$84=\$1284$

Try it Now 1

A TV originally priced at $799 is on sale for 30% off. There is then a 9.2% sales tax. Find the price after including the discount and sales tax.

Answer

The sale price is $\$799(0.70)=\$559.30$ After tax, the price is $\$559.30(1.092)=\$610.76$

Example 5

The value of a car dropped from $7400 to $6800 over the last year. What percent decrease is this?

Solution

To compute the percent change, we first need to find the dollar value change: $\$6800-\$7400=-\$600$ Often we will take the absolute value of this amount, which is called the **absolute change**: $|-600|=600$.

Since we are computing the decrease relative to the starting value, we compute this percent out of $\$7400$:

$\frac{600}{7400}=0.081=8.1\%$ decrease. This is called a **relative change**.

Absolute and Relative Change

Given two quantities,

Absolute change = $|\text{ending quantity } - \text{ starting quantity}|$

Relative change: $\frac{\text{absolute change}}{\text{starting quantity}}$

Absolute change has the same units as the original quantity.

Relative change gives a percent change.

The starting quantity is called the **base** of the percent change.

The base of a percent is very important. For example, while Nixon was president, it was argued that marijuana was a "gateway" drug, claiming that 80% of marijuana smokers went on to use harder drugs like cocaine. The problem is, this isn't true. The true claim is that 80% of harder drug users first smoked marijuana. The difference is one of base: 80% of marijuana smokers using hard drugs, vs. 80% of hard drug users having smoked marijuana. These numbers are not equivalent. As it turns out, only one in 2,400 marijuana users actually go on to use harder drugs[2].

Example 6

There are about 75 QFC supermarkets in the U.S. Albertsons has about 215 stores. Compare the size of the two companies.

Solution

When we make comparisons, we must ask first whether an absolute or relative comparison. The absolute difference is $215-75=140$. From this, we could say "Albertsons has 140 more stores than QFC." However, if you wrote this in an article or paper, that number does not mean much. The relative difference may be more meaningful. There are two different relative changes we could calculate, depending on which store we use as the base:

Using QFC as the base, $\frac{140}{75}=1.867$.

This tells us Albertsons is 186.7% larger than QFC.

Using Albertsons as the base, $\frac{140}{215}=0.651$.

This tells us QFC is 65.1% smaller than Albertsons.

Notice both of these are showing percent *differences*. We could also calculate the size of Albertsons relative to QFC: , which tells us Albertsons is 2.867 times the size of QFC. Likewise, we could calculate the size of QFC relative to Albertsons: , which tells us that QFC is 34.9% of the size of Albertsons.

Example 7

Suppose a stock drops in value by 60% one week, then increases in value the next week by 75%. Is the value higher or lower than where it started?

Solution

To answer this question, suppose the value started at $100. After one week, the value dropped by 60%:

$$\$100 - \$100(0.60) = \$100 - \$60 = \$40.$$

In the next week, notice that base of the percent has changed to the new value, $40. Computing the 75% increase:

$$\$40 + \$40(0.75) = \$40 + \$30 = \$70.$$

In the end, the stock is still $30 lower, or $\frac{\$30}{\$100} = 30\%$ lower, valued than it started.

Try it Now 2

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[3]. Calculate the absolute and relative increase for 2001-2002 and 2005-2006. Which year saw a larger increase in federal debt?

Answer

2001-2002: Absolute change: $0.43 trillion. Relative change: 7.45%

2005-2006: Absolute change: $0.54 trillion. Relative change: 6.83%

2005-2006 saw a larger absolute increase, but a smaller relative increase.

Example 8

A Seattle Times article on high school graduation rates reported “The number of schools graduating 60 percent or fewer students in four years – sometimes referred to as “dropout factories” – decreased by 17 during that time period. The number of kids attending schools with such low graduation rates was cut in half.”

a) Is the “decrease by 17” number a useful comparison?

b) Considering the last sentence, can we conclude that the number of “dropout factories” was originally 34?

Solution

a) This number is hard to evaluate, since we have no basis for judging whether this is a larger or small change. If the number of “dropout factories” dropped from 20 to 3, that’d be a very significant change, but if the number dropped from 217 to 200, that’d be less of an improvement.

b) The last sentence provides relative change which helps put the first sentence in perspective. We can estimate that the number of “dropout factories” was probably previously around 34. However, it’s possible that students simply moved schools rather than the school improving, so that estimate might not be fully accurate.

Example 9

In the 2004 vice-presidential debates, Edwards's claimed that US forces have suffered "90% of the coalition casualties" in Iraq. Cheney disputed this, saying that in fact Iraqi security forces and coalition allies "have taken almost 50 percent" of the casualties. Who is correct?

Solution

Without more information, it is hard for us to judge who is correct, but we can easily conclude that these two percents are talking about different things, so one does not necessarily contradict the other. Edward's claim was a percent with coalition forces as the base of the percent, while Cheney's claim was a percent with both coalition and Iraqi security forces as the base of the percent. It turns out both statistics are in fact fairly accurate.

? Try it Now 3

In the 2012 presidential elections, one candidate argued that "the president's plan will cut $716 billion from Medicare, leading to fewer services for seniors," while the other candidate rebuts that "our plan does not cut current spending and actually expands benefits for seniors, while implementing cost saving measures." Are these claims in conflict, in agreement, or not comparable because they're talking about different things?

Answer

Without more information, it is hard to judge these arguments. This is compounded by the complexity of Medicare. As it turns out, the $716 billion is not a cut in current spending, but a cut in future increases in spending, largely reducing future growth in health care payments. In this case, at least the numerical claims in both statements could be considered at least partially true. Here is one source of more information if you're interested: http://factcheck.org/2012/08/a-campaign-full-of-mediscare/

We'll wrap up our review of percents with a couple cautions. First, when talking about a change of quantities that are already measured in percents, we have to be careful in how we describe the change.

✓ Example 10

A politician's support increases from 40% of voters to 50% of voters. Describe the change.

Solution

We could describe this using an absolute change: $|50\% - 40\%| = 10\%$. Notice that since the original quantities were percents, this change also has the units of percent. In this case, it is best to describe this as an increase of 10 **percentage points**.

In contrast, we could compute the percent change: $\frac{10\%}{40\%} = 0.25 = 25\%$ increase. This is the relative change, and we'd say the politician's support has increased by 25%.

Lastly, a caution against averaging percents.

✓ Example 11

A basketball player scores on 40% of 2-point field goal attempts, and on 30% of 3-point of field goal attempts. Find the player's overall field goal percentage.

Solution

It is very tempting to average these values, and claim the overall average is 35%, but this is likely not correct, since most players make many more 2-point attempts than 3-point attempts. We don't actually have enough information to answer the question. Suppose the player attempted 200 2-point field goals and 100 3-point field goals. Then they made 200(0.40) = 80 2-point shots and $100(0.30) = 30$ 3-point shots. Overall, they made 110 shots out of 300, for a $\frac{110}{300} = 0.367 = 36.7\%$ overall field goal percentage.

---

[1] www.factcheck.org/cheney_edwards_mangle_facts.html

[2] http://tvtropes.org/pmwiki/pmwiki.php/Main/LiesDamnedLiesAndStatistics

[3] www.whitehouse.gov/sites/defa...s/hist07z1.xls

- **1.2: Percents** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

1.3: Proportions and Rates

If you wanted to power the city of Seattle using wind power, how many windmills would you need to install? Questions like these can be answered using rates and proportions.

Rates

A rate is the ratio (fraction) of two quantities.

A **unit rate** is a rate with a denominator of one.

Example 12

Your car can drive 300 miles on a tank of 15 gallons. Express this as a rate.

Solution

Expressed as a rate, $\frac{300\text{ miles}}{15\text{ gallons}}$. We can divide to find a unit rate: $\frac{20\text{ miles}}{1\text{gallon}}$, which we could also write as $20\frac{\text{miles}}{\text{gallon}}$, or just 20 miles per gallon.

Proportion Equation

A proportion equation is an equation showing the equivalence of two rates or ratios.

Example 13

Solve the proportion $\frac{5}{3}=\frac{x}{6}$ for the unknown value x.

Solution

This proportion is asking us to find a fraction with denominator 6 that is equivalent to the fraction $\frac{5}{3}$. We can solve this by multiplying both sides of the equation by 6, giving $x=\frac{5}{3}\cdot 6=10$.

Example 14

A map scale indicates that $\frac{1}{2}$ inch on the map corresponds with 3 real miles. How many miles apart are two cities that are $2\frac{1}{4}$ inches apart on the map?

Solution

We can set up a proportion by setting equal two $\frac{\text{map inches}}{\text{real miles}}$ rates, and introducing a variable, x, to represent the unknown quantity – the mile distance between the cities.

$$\frac{\frac{1}{2}\text{ map inch}}{3\text{ miles}}=\frac{2\frac{1}{4}\text{ map inches}}{x\text{ miles}} \qquad \text{Multiply both sides by } x \text{ and rewriting the mixed number}$$

$$\frac{\frac{1}{2}}{3}\cdot x=\frac{9}{4} \qquad \text{Multiply both sides by 3}$$

$$\frac{1}{2}x=\frac{27}{4} \qquad \text{Multiply both sides by 2 (or divide by } \tfrac{1}{2})$$

$$x=\frac{27}{2}=13\frac{1}{2}\text{ miles}$$

Many proportion problems can also be solved using **dimensional analysis**, the process of multiplying a quantity by rates to change the units.

Example 15

Your car can drive 300 miles on a tank of 15 gallons. How far can it drive on 40 gallons?

Solution

We could certainly answer this question using a proportion: $ $\frac{300 \text{ miles}}{15 \text{ gallons}} = \frac{x \text{ miles}}{40 \text{ gallons}}$ $.

However, we earlier found that 300 miles on 15 gallons gives a rate of 20 miles per gallon. If we multiply the given 40 gallon quantity by this rate, the *gallons* unit “cancels” and we’re left with a number of miles:

$$40\text{gallons} \cdot \frac{20 \text{ miles}}{\text{gallon}} = \frac{40 \text{ gallons}}{1} \cdot \frac{20 \text{ miles}}{\text{gallon}} = 800 \text{ miles}$$

Notice if instead we were asked “how many gallons are needed to drive 50 miles?” we could answer this question by inverting the 20 mile per gallon rate so that the *miles* unit cancels and we’re left with gallons:

$$50\text{miles} \cdot \frac{1\text{gallon}}{20 \text{ miles}} = \frac{50 \text{ miles}}{1} \cdot \frac{1 \text{ gallon}}{20 \text{ miles}} = \frac{50 \text{ gallons}}{20} = 2.5 \text{ gallons}$$

Dimensional analysis can also be used to do unit conversions. Here are some unit conversions for reference.

Unit Conversions

Length

$1 \text{ foot (ft)} = 12 \text{ inches (in)}$ $\quad$ $1 \text{ yard (yd)} = 3 \text{ feet (ft)}$

$1 \text{ mile} = 5,280 \text{ feet}$

$1000 \text{ millimeters } mm = 1 \text{ meter (m)}$ $\quad$ $100 \text{ centimeters (cm)} = 1 \text{ meter}$

$1000 \text{ meters (m)} = 1 \text{ kilometer (km)}$ $\quad$ $2.54 \text{ centimeters (cm)} = 1 \text{ inch}$

Weight and Mass

$1 \text{ pound (lb)} = 16 \text{ ounces (oz)}$ $\quad$ $1 \text{ ton} = 2000 \text{ pounds}$

$1000 \text{ milligrams (mg)} = 1 \text{ gram (g)}$ $\quad$ $1000 \text{ grams} = 1 \text{ kilogram (kg)}$

$1 \text{ kilogram} = 2.2 \text{ pounds (on earth)}$

Capacity

$1 \text{ cup} = 8 \text{ fluid ounces (fl oz)}^*$ $\quad$ $1 \text{ pint} = 2 \text{ cups}$

$1 \text{ quart} = 2 \text{ pints} = 4 \text{ cups}$ $\quad$ $1 \text{ gallon} = 4 \text{ quarts} = 16 \text{ cups}$

$1000 \text{ milliliters (ml)} = 1 \text{ liter (L)}$

*Fluid ounces are a capacity measurement for liquids. 1 fluid ounce ≈ 1 ounce (weight) for water only.

Example 16

A bicycle is traveling at 15 miles per hour. How many feet will it cover in 20 seconds?

Solution

To answer this question, we need to convert 20 seconds into feet. If we know the speed of the bicycle in feet per second, this question would be simpler. Since we don’t, we will need to do additional unit conversions. We will need to know that 5280 ft = 1 mile. We might start by converting the 20 seconds into hours:

$$20 \text{ seconds} \cdot \frac{1 \text{ minute}}{60 \text{ seconds}} \cdot \frac{1 \text{ hour}}{60 \text{ minutes}} = \frac{1}{180} \text{ hour} \qquad \text{Now we can multiply by the 15 miles/hr}$$

$$\frac{1}{180} \text{ hour} \cdot \frac{15 \text{ miles}}{1\text{hour}} = \frac{1}{12} \text{ mile} \qquad \text{Now we can convert to feet}$$

$$\frac{1}{12} \text{ mile} \cdot \frac{5280 \text{ feet}}{1 \text{ mile}} = 440 \text{ feet}$$

We could have also done this entire calculation in one long set of products:

$$20 \text{ seconds} \cdot \frac{1 \text{ minute}}{60 \text{ seconds}} \cdot \frac{1 \text{ hour}}{60 \text{ minutes}} \cdot \frac{15 \text{ miles}}{1 \text{ hour}} \cdot \frac{5280 \text{ feet}}{1 \text{ mile}} = 440 \text{ feet}$$

Try it Now 4

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?

Answer

$$18 \text{ inches} \cdot \frac{1 \text{ foot}}{12 \text{ inches}} \cdot \frac{19.8 \text{ pounds}}{1000 \text{ feet}} \cdot \frac{16 \text{ ounces}}{1 \text{ pound}} \approx 0.475 \text{ ounces}$$

Notice that with the miles per gallon example, if we double the miles driven, we double the gas used. Likewise, with the map distance example, if the map distance doubles, the real-life distance doubles. This is a key feature of proportional relationships, and one we must confirm before assuming two things are related proportionally.

Example 17

Suppose you're tiling the floor of a 10 ft by 10 ft room, and find that 100 tiles will be needed. How many tiles will be needed to tile the floor of a 20 ft by 20 ft room?

Solution

In this case, while the width the room has doubled, the area has quadrupled. Since the number of tiles needed corresponds with the area of the floor, not the width, 400 tiles will be needed. We could find this using a proportion based on the areas of the rooms:

$$\frac{100 \text{ tiles}}{100\text{ft}^2} = \frac{n \text{ tiles}}{400\text{ft}^2}$$

Other quantities just don't scale proportionally at all.

Example 18

Suppose a small company spends $1000 on an advertising campaign, and gains 100 new customers from it. How many new customers should they expect if they spend $10,000?

Solution

While it is tempting to say that they will gain 1000 new customers, it is likely that additional advertising will be less effective than the initial advertising. For example, if the company is a hot tub store, there are likely only a fixed number of people interested in buying a hot tub, so there might not even be 1000 people in the town who would be potential customers.

Sometimes when working with rates, proportions, and percents, the process can be made more challenging by the magnitude of the numbers involved. Sometimes, large numbers are just difficult to comprehend.

Example 19

Compare the 2010 U.S. military budget of $683.7 billion to other quantities.

Here we have a very large number, about $683,700,000,000 written out. Of course, imagining a billion dollars is very difficult, so it can help to compare it to other quantities.

If that amount of money was used to pay the salaries of the 1.4 million Walmart employees in the U.S., each would earn over $488,000.

There are about 300 million people in the U.S. The military budget is about $2,200 per person.

If you were to put $683.7 billion in $100 bills, and count out 1 per second, it would take 216 years to finish counting it.

Example 20

Compare the electricity consumption per capita in China to the rate in Japan.

To address this question, we will first need data. From the CIA[1] website we can find the electricity consumption in 2011 for China was 4,693,000,000,000 KWH (kilowatt-hours), or 4.693 trillion KWH, while the consumption for Japan was 859,700,000,000, or 859.7 billion KWH. To find the rate per capita (per person), we will also need the population of the two countries. From the World Bank[2], we can find the population of China is 1,344,130,000, or 1.344 billion, and the population of Japan is 127,817,277, or 127.8 million.

Solution

Computing the consumption per capita for each country:

China: $\frac{4{,}693{,}000{,}000{,}000 \text{KWH}}{1{,}344{,}130{,}000 \text{ people}} \approx 3491.5$ KWH per person

Japan: $\frac{859{,}700{,}000{,}000 \text{KWH}}{127{,}817{,}277 \text{ people}} \approx 6726$ KWH per person

While China uses more than 5 times the electricity of Japan overall, because the population of Japan is so much smaller, it turns out Japan uses almost twice the electricity per person compared to China.

---

[1] www.cia.gov/library/publicat.../2042rank.html

[2] http://data.worldbank.org/indicator/SP.POP.TOT

1.4: Geometry

Geometric shapes, as well as area and volumes, can often be important in problem solving.

Example 21

You are curious how tall a tree is, but don't have any way to climb it. Describe a method for determining the height.

There are several approaches we could take. We'll use one based on triangles, which requires that it's a sunny day. Suppose the tree is casting a shadow, say 15 ft long. I can then have a friend help me measure my own shadow. Suppose I am 6 ft tall, and cast a 1.5 ft shadow. Since the triangle formed by the tree and its shadow has the same angles as the triangle formed by me and my shadow, these triangles are called **similar triangles** and their sides will scale proportionally. In other words, the ratio of height to width will be the same in both triangles. Using this, we can find the height of the tree, which we'll denote by h:

Solution

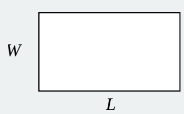
$$\frac{6\text{ft tall}}{1.5\text{ft shadow}} = \frac{h\text{ft tall}}{15\text{ft shadow}}$$

Multiplying both sides by 15, we get $h = 60$. The tree is about 60 ft tall.

It may be helpful to recall some formulas for areas and volumes of a few basic shapes.

Areas

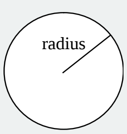
Rectangle



Area: $L \cdot W$

Perimeter: $2L + 2W$

Circle, radius r

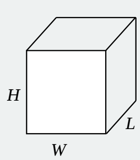


Area: πr^2

Circumference = $2\pi r$

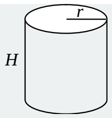
Volumes

Rectangular Box



Volume: $L \cdot W \cdot H$

Cylinder



Volume: $\pi r^2 H$

Example 22

If a 12 inch diameter pizza requires 10 ounces of dough, how much dough is needed for a 16 inch pizza?

Solution

To answer this question, we need to consider how the weight of the dough will scale. The weight will be based on the volume of the dough. However, since both pizzas will be about the same thickness, the weight will scale with the area of the top of the pizza. We can find the area of each pizza using the formula for area of a circle, $A = \pi r^2$:

A 12” pizza has radius 6 inches, so the area will be $\pi 6^2$ = about 113 square inches.

A 16” pizza has radius 8 inches, so the area will be $\pi 8^2$ = about 201 square inches.

Notice that if both pizzas were 1 inch thick, the volumes would be 113 in^3 and 201 in^3 respectively, which are at the same ratio as the areas. As mentioned earlier, since the thickness is the same for both pizzas, we can safely ignore it.

We can now set up a proportion to find the weight of the dough for a 16” pizza:

$\frac{10 \text{ ounces}}{113 \text{in}^2} = \frac{x \text{ ounces}}{201 \text{in}^2}$ Multiply both sides by 201

$x = 201 \cdot \frac{10}{113} =$ about 17.8 ounces of dough for a 16” pizza.

It is interesting to note that while the diameter is $\frac{16}{12} = 1.33$ times larger, the dough required, which scales with area, is $1.33^2 = 1.78$ times larger.

Example 23

A company makes regular and jumbo marshmallows. The regular marshmallow has 25 calories. How many calories will the jumbo marshmallow have?

Photo courtesy Christopher Danielson

We would expect the calories to scale with volume. Since the marshmallows have cylindrical shapes, we can use that formula to find the volume. From the grid in the image, we can estimate the radius and height of each marshmallow.

Solution

The regular marshmallow appears to have a diameter of about 3.5 units, giving a radius of 1.75 units, and a height of about 3.5 units. The volume is about $\pi(1.75)^2(3.5) = 33.7 \text{ units}^3$.

The jumbo marshmallow appears to have a diameter of about 5.5 units, giving a radius of 2.75 units, and a height of about 5 units. The volume is about $\pi(2.75)^2(5) = 118.8 \text{ units}^3$.

We could now set up a proportion, or use rates. The regular marshmallow has 25 calories for 33.7 cubic units of volume. The jumbo marshmallow will have:

$118.8 \text{ units}^3 \cdot \frac{25 \text{ calories}}{33.7 \text{ units}^3} = 88.1 \text{ calories}$

It is interesting to note that while the diameter and height are about 1.5 times larger for the jumbo marshmallow, the volume and calories are about $1.5^3 = 3.375$ times larger.

Try it Now 5

A website says that you'll need 48 fifty-pound bags of sand to fill a sandbox that measure 8ft by 8ft by 1ft. How many bags would you need for a sandbox 6ft by 4ft by 1ft?

Answer

The original sandbox has volume 64ft^3. The smaller sandbox has volume 24ft^3.

$\frac{48\text{bags}}{64\text{ft}^3} = \frac{x\text{bags}}{24\text{ft}^3}$ results in $x = 18$ bags.

1.5: Problem Solving and Estimating

Finally, we will bring together the mathematical tools we've reviewed, and use them to approach more complex problems. In many problems, it is tempting to take the given information, plug it into whatever formulas you have handy, and hope that the result is what you were supposed to find. Chances are, this approach has served you well in other math classes.

This approach does not work well with real life problems. Instead, problem solving is best approached by first starting at the end: identifying exactly what you are looking for. From there, you then work backwards, asking "what information and procedures will I need to find this?" Very few interesting questions can be answered in one mathematical step; often times you will need to chain together a solution pathway, a series of steps that will allow you to answer the question.

Problem Solving Process

1. Identify the question you're trying to answer.
2. Work backwards, identifying the information you will need and the relationships you will use to answer that question.
3. Continue working backwards, creating a solution pathway.
4. If you are missing necessary information, look it up or estimate it. If you have unnecessary information, ignore it.
5. Solve the problem, following your solution pathway.

In most problems we work, we will be approximating a solution, because we will not have perfect information. We will begin with a few examples where we will be able to approximate the solution using basic knowledge from our lives.

Example 24

How many times does your heart beat in a year?

Solution

This question is asking for the rate of heart beats per year. Since a year is a long time to measure heart beats for, if we knew the rate of heart beats per minute, we could scale that quantity up to a year. So the information we need to answer this question is heart beats per minute. This is something you can easily measure by counting your pulse while watching a clock for a minute.

Suppose you count 80 beats in a minute. To convert this beats per year:

$\frac{80 \text{ beats}}{1 \text{ minute}} \cdot \frac{60 \text{ minutes}}{1 \text{ hour}} \cdot \frac{24 \text{ hours}}{1 \text{ day}} \cdot \frac{365 \text{ days}}{1 \text{ year}} = 42,048,000 \text{ beats per year}$

Example 25

How thick is a single sheet of paper? How much does it weigh?

Solution

While you might have a sheet of paper handy, trying to measure it would be tricky. Instead we might imagine a stack of paper, and then scale the thickness and weight to a single sheet. If you've ever bought paper for a printer or copier, you probably bought a ream, which contains 500 sheets. We could estimate that a ream of paper is about 2 inches thick and weighs about 5 pounds. Scaling these down,

$\frac{2 \text{ inches}}{\text{ream}} \cdot \frac{1 \text{ ream}}{500 \text{ pages}} = 0.004 \text{ inches per sheet}$

$\frac{5 \text{ pounds}}{\text{ream}} \cdot \frac{1 \text{ ream}}{500 \text{ pages}} = 0.01 \text{ pounds per sheet, or 0.16 ounces per sheet.}$

Example 26

A recipe for zucchini muffins states that it yields 12 muffins, with 250 calories per muffin. You instead decide to make mini-muffins, and the recipe yields 20 muffins. If you eat 4, how many calories will you consume?

Solution

There are several possible solution pathways to answer this question. We will explore one.

To answer the question of how many calories 4 mini-muffins will contain, we would want to know the number of calories in each mini-muffin. To find the calories in each mini-muffin, we could first find the total calories for the entire recipe, then divide it by the number of mini-muffins produced. To find the total calories for the recipe, we could multiply the calories per standard muffin by the number per muffin. Notice that this produces a multi-step solution pathway. It is often easier to solve a problem in small steps, rather than trying to find a way to jump directly from the given information to the solution.

We can now execute our plan:

$12 \text{muffins\$} \cdot \frac{250 \text{ calories}}{\text{muffin}} = 3000\$$ calories for the whole recipe

$\frac{3000 \text{ calories}}{20 \text{ mini} - \text{muffins}}$ gives 150 calories per mini-muffin

$4 \text{ mini muffins} \cdot \frac{150 \text{ calories}}{\text{mini - muffin}}$ totals 600 calories consumed.

✔ Example 27

You need to replace the boards on your deck. About how much will the materials cost?

Solution

There are two approaches we could take to this problem: 1) estimate the number of boards we will need and find the cost per board, or 2) estimate the area of the deck and find the approximate cost per square foot for deck boards. We will take the latter approach.

For this solution pathway, we will be able to answer the question if we know the cost per square foot for decking boards and the square footage of the deck. To find the cost per square foot for decking boards, we could compute the area of a single board, and divide it into the cost for that board. We can compute the square footage of the deck using geometric formulas. So first we need information: the dimensions of the deck, and the cost and dimensions of a single deck board.

Suppose that measuring the deck, it is rectangular, measuring 16 ft by 24 ft, for a total area of 384ft^2.

From a visit to the local home store, you find that an 8 foot by 4 inch cedar deck board costs about \$7.50. The area of this board, doing the necessary conversion from inches to feet, is:

$8 \text{ feet } \cdot 4 \text{ inches} \cdot \frac{1 \text{ foot}}{12 \text{ inches}} = 2.667 \text{ft}^2$. The cost per square foot is then

$\frac{\$7.50}{2.667 \text{ft}^2} = \$2.8125 \text{ per ft}^2.$

This will allow us to estimate the material cost for the whole 384ft^2 deck

$\$384 \text{ft}^2 \cdot \frac{\$2.8125}{\text{ft}^2} = \$1080$ total cost.

Of course, this cost estimate assumes that there is no waste, which is rarely the case. It is common to add at least 10% to the cost estimate to account for waste.

✔ Example 28

Is it worth buying a Hyundai Sonata hybrid instead the regular Hyundai Sonata?

Solution

To make this decision, we must first decide what our basis for comparison will be. For the purposes of this example, we'll focus on fuel and purchase costs, but environmental impacts and maintenance costs are other factors a buyer might consider.

It might be interesting to compare the cost of gas to run both cars for a year. To determine this, we will need to know the miles per gallon both cars get, as well as the number of miles we expect to drive in a year. From that information, we can find the number of gallons required from a year. Using the price of gas per gallon, we can find the running cost.

From Hyundai's website, the 2013 Sonata will get 24 miles per gallon (mpg) in the city, and 35 mpg on the highway. The hybrid will get 35 mpg in the city, and 40 mpg on the highway.

An average driver drives about 12,000 miles a year. Suppose that you expect to drive about 75% of that in the city, so 9,000 city miles a year, and 3,000 highway miles a year.

We can then find the number of gallons each car would require for the year.

Sonata:

$9000 \text{ city miles} \cdot \frac{1 \text{ gallon}}{24 \text{ city miles}} + 3000 \text{ hightway miles.} \frac{1 \text{ gallon}}{35 \text{ highway miles}} = 460.7 \text{ gallons}$

Hybrid:

$9000 \text{ city miles} \cdot \frac{1 \text{ gallon}}{35 \text{ city miles}} + 3000 \text{ hightway miles.} \frac{1 \text{ gallon}}{40 \text{ highway miles}} = 332.1 \text{ gallons}$

If gas in your area averages about $3.50 per gallon, we can use that to find the running cost:

Sonata: $460.7 \text{ gallons} \cdot \frac{\$3.50}{\text{gallon}} = \$1612.45$

Hybrid: $332.1 \text{ gallons} \cdot \frac{\$3.50}{\text{gallon}} = \$1162.35$

The hybrid will save $450.10 a year. The gas costs for the hybrid are about $\frac{\$450.10}{\$1612.45} = 0.279 = 27.9\%$ lower than the costs for the standard Sonata.

While both the absolute and relative comparisons are useful here, they still make it hard to answer the original question, since “is it worth it” implies there is some tradeoff for the gas savings. Indeed, the hybrid Sonata costs about $25,850, compared to the base model for the regular Sonata, at $20,895.

To better answer the “is it worth it” question, we might explore how long it will take the gas savings to make up for the additional initial cost. The hybrid costs $4965 more. With gas savings of $451.10 a year, it will take about 11 years for the gas savings to make up for the higher initial costs.

We can conclude that if you expect to own the car 11 years, the hybrid is indeed worth it. If you plan to own the car for less than 11 years, it may still be worth it, since the resale value of the hybrid may be higher, or for other non-monetary reasons. This is a case where math can help guide your decision, but it can’t make it for you.

? Try it Now 6

If traveling from Seattle, WA to Spokane WA for a three-day conference, does it make more sense to drive or fly?

Answer

There is not enough information provided to answer the question, so we will have to make some assumptions, and look up some values.

Assumptions:

a) We own a car. Suppose it gets 24 miles to the gallon. We will only consider gas cost.

b) We will not need to rent a car in Spokane, but will need to get a taxi from the airport to the conference hotel downtown and back.

c) We can get someone to drop us off at the airport, so we don’t need to consider airport parking.

d) We will not consider whether we will lose money by having to take time off work to drive.

Values looked up (your values may be different)

a) Flight cost: $\$184$

b) Taxi cost: $\$25$ each way (estimate, according to hotel website)

c) Driving distance: 280 miles each way

d) Gas cost: $\$3.79$ a gallon

Cost for flying: $\$184 \text{ flight cost } + \$50 \text{ in taxi fares } = \$234$.

Cost for driving: 560 miles round trip will require 23.3 gallons of gas, costing $\$88.31$

Based on these assumptions, driving is cheaper. However, our assumption that we only include gas cost may not be a good one. Tax law allows you deduct $\$0.55$ (in 2012) for each mile driven, a value that accounts for gas as well as a portion of the car cost, insurance, maintenance, etc. Based on this number, the cost of driving would be $\$319$.

1.6: Exercises

1. Out of 230 racers who started the marathon, 212 completed the race, 14 gave up, and 4 were disqualified. What percentage did not complete the marathon?

2. Patrick left an $8 tip on a $50 restaurant bill. What percent tip is that?

3. Ireland has a 23% VAT (value-added tax, similar to a sales tax). How much will the VAT be on a purchase of a €250 item?

4. Employees in 2012 paid 4.2% of their gross wages towards social security (FICA tax), while employers paid another 6.2%. How much will someone earning $45,000 a year pay towards social security out of their gross wages?

5. A project on Kickstarter.com was aiming to raise $15,000 for a precision coffee press. They ended up with 714 supporters, raising 557% of their goal. How much did they raise?

6. Another project on Kickstarter for an iPad stylus raised 1,253% of their goal, raising a total of $313,490 from 7,511 supporters. What was their original goal?

7. The population of a town increased from 3,250 in 2008 to 4,300 in 2010. Find the absolute and relative (percent) increase.

8. The number of CDs sold in 2010 was 114 million, down from 147 million the previous year[1]. Find the absolute and relative (percent) decrease.

9. A company wants to decrease their energy use by 15%.

   a. If their electric bill is currently $2,200 a month, what will their bill be if they're successful?

   b. If their next bill is $1,700 a month, were they successful? Why or why not?

10. A store is hoping an advertising campaign will increase their number of customers by 30%. They currently have about 80 customers a day.

    a. How many customers will they have if their campaign is successful?

    b. If they increase to 120 customers a day, were they successful? Why or why not?

11. An article reports "attendance dropped 6% this year, to 300." What was the attendance before the drop?

12. An article reports "sales have grown by 30% this year, to $200 million." What were sales before the growth?

13. The Walden University had 47,456 students in 2010, while Kaplan University had 77,966 students. Complete the following statements:

    a. Kaplan's enrollment was ___% larger than Walden's.

    b. Walden's enrollment was ___% smaller than Kaplan's.

    c. Walden's enrollment was ___% of Kaplan's.

14. In the 2012 Olympics, Usain Bolt ran the 100m dash in 9.63 seconds. Jim Hines won the 1968 Olympic gold with a time of 9.95 seconds.

    a. Bolt's time was ___% faster than Hines'.

    b. Hine' time was ___% slower than Bolt's.

    c. Hine' time was ___% of Bolt's.

15. A store has clearance items that have been marked down by 60%. They are having a sale, advertising an additional 30% off clearance items. What percent of the original price do you end up paying?

16. Which is better: having a stock that goes up 30% on Monday than drops 30% on Tuesday, or a stock that drops 30% on Monday and goes up 30% on Tuesday? In each case, what is the net percent gain or loss?

17. Are these two claims equivalent, in conflict, or not comparable because they're talking about different things?

    a. "16.3% of Americans are without health insurance"[2]

    b. "only 55.9% of adults receive employer provided health insurance"[3]

18. Are these two claims equivalent, in conflict, or not comparable because they’re talking about different things?

a. “We mark up the wholesale price by 33% to come up with the retail price”

b. “The store has a 25% profit margin”

19. Are these two claims equivalent, in conflict, or not comparable because they’re talking about different things?

a. “Every year since 1950, the number of American children gunned down has doubled.”

b. “The number of child gunshot deaths has doubled from 1950 to 1994.”

20. Are these two claims equivalent, in conflict, or not comparable because they’re talking about different things?[4]

a. “75 percent of the federal health care law’s taxes would be paid by those earning less than $120,000 a year”

b. “76 percent of those who would pay the penalty [health care law’s taxes] for not having insurance in 2016 would earn under $120,000”

21. Are these two claims equivalent, in conflict, or not comparable because they’re talking about different things?

a. “The school levy is only a 0.1% increase of the property tax rate.”

b. “This new levy is a 12% tax hike, raising our total rate to $9.33 per $1000 of value.”

22. Are the values compared in this statement comparable or not comparable? “Guns have murdered more Americans here at home in recent years than have died on the battlefields of Iraq and Afghanistan. In support of the two wars, more than 6,500 American soldiers have lost their lives. During the same period, however, guns have been used to murder about 100,000 people on American soil”[5]

23. A high school currently has a 30% dropout rate. They’ve been tasked to decrease that rate by 20%. Find the equivalent percentage point drop.

24. A politician’s support grew from 42% by 3 percentage points to 45%. What percent (relative) change is this?

25. Marcy has a 70% average in her class going into the final exam. She says "I need to get a 100% on this final so I can raise my score to 85%." Is she correct?

26. Suppose you have one quart of water/juice mix that is 50% juice, and you add 2 quarts of juice. What percent juice is the final mix?

27. Find a unit rate: You bought 10 pounds of potatoes for $4.

28. Find a unit rate: Joel ran 1500 meters in 4 minutes, 45 seconds.

29. Solve: $\frac{2}{5} = \frac{6}{x}$.

30. Solve: $\frac{n}{5} = \frac{16}{20}$.

31. A crepe recipe calls for 2 eggs, 1 cup of flour, and 1 cup of milk. How much flour would you need if you use 5 eggs?

32. An 8ft length of 4 inch wide crown molding costs $14. How much will it cost to buy 40ft of crown molding?

33. Four 3-megawatt wind turbines can supply enough electricity to power 3000 homes. How many turbines would be required to power 55,000 homes?

34. A highway had a landslide, where 3,000 cubic yards of material fell on the road, requiring 200 dump truck loads to clear. On another highway, a slide left 40,000 cubic yards on the road. How many dump truck loads would be needed to clear this slide?

35. Convert 8 feet to inches.

36. Convert 6 kilograms to grams.

37. A wire costs $2 per meter. How much will 3 kilometers of wire cost?

38. Sugar contains 15 calories per teaspoon. How many calories are in 1 cup of sugar?

39. A car is driving at 100 kilometers per hour. How far does it travel in 2 seconds?

40. A chain weighs 10 pounds per foot. How many ounces will 4 inches weigh?

41. The table below gives data on three movies. Gross earnings is the amount of money the movie brings in. Compare the net earnings (money made after expenses) for the three movies.[6]

Movie Earnings

Movie	Release Date	Budget	Gross Earnings
Saw	10/29/2004	$1,200,000	$103,096,345
Titanic	10/29/2004	$200,000,000	$1,842,879,955
Jurassic Park	6/11/1993	$63,000,000	$923,863,984

42. For the movies in the previous problem, which provided the best return on investment?

43. The population of the U.S. is about 309,975,000, covering a land area of 3,717,000 square miles. The population of India is about 1,184,639,000, covering a land area of 1,269,000 square miles. Compare the population densities of the two countries.

44. The GDP (Gross Domestic Product) of China was $5,739 billion in 2010, and the GDP of Sweden was $435 billion. The population of China is about 1,347 million, while the population of Sweden is about 9.5 million. Compare the GDP per capita of the two countries.

45. In June 2012, Twitter was reporting 400 million tweets per day. Each tweet can consist of up to 140 characters (letter, numbers, etc.). Create a comparison to help understand the amount of tweets in a year by imagining each character was a drop of water and comparing to filling something up.

46. The photo sharing site Flickr had 2.7 billion photos in June 2012. Create a comparison to understand this number by assuming each picture is about 2 megabytes in size, and comparing to the data stored on other media like DVDs, iPods, or flash drives.

47. Your chocolate milk mix says to use 4 scoops of mix for 2 cups of milk. After pouring in the milk, you start adding the mix, but get distracted and accidentally put in 5 scoops of mix. How can you adjust the mix if:

a. There is still room in the cup?

b. The cup is already full?

48. A recipe for sabayon calls for 2 egg yolks, 3 tablespoons of sugar, and $\frac{1}{4}$ cup of white wine. After cracking the eggs, you start measuring the sugar, but accidentally put in 4 tablespoons of sugar. How can you compensate?

49. The Deepwater Horizon oil spill resulted in 4.9 million barrels of oil spilling into the Gulf of Mexico. Each barrel of oil can be processed into about 19 gallons of gasoline. How many cars could this have fueled for a year? Assume an average car gets 20 miles to the gallon, and drives about 12,000 miles in a year.

50. The store is selling lemons at 2 for $1. Each yields about 2 tablespoons of juice. How much will it cost to buy enough lemons to make a 9-inch lemon pie requiring $\frac{1}{2}$ cup of lemon juice?

51. A piece of paper can be made into a cylinder in two ways: by joining the short sides together, or by joining the long sides together[7]. Which cylinder would hold more? How much more?

52. Which of these glasses contains more liquid? How much more?

In the next 4 questions, estimate the values by making reasonable approximations for unknown values, or by doing some research to find reasonable values.

53. Estimate how many gallons of water you drink in a year.

54. Estimate how many times you blink in a day.

55. How much does the water in a 6-person hot tub weigh?

56. How many gallons of paint would be needed to paint a two-story house 40 ft long and 30 ft wide?

57. During the landing of the Mars Science Laboratory *Curiosity*, it was reported that the signal from the rover would take 14 minutes to reach earth. Radio signals travel at the speed of light, about 186,000 miles per second. How far was Mars from Earth when *Curiosity* landed?

58. It is estimated that a driver takes, on average, 1.5 seconds from seeing an obstacle to reacting by applying the brake or swerving. How far will a car traveling at 60 miles per hour travel (in feet) before the driver reacts to an obstacle?

59. The flash of lightning travels at the speed of light, which is about 186,000 miles per second. The sound of lightning (thunder) travels at the speed of sound, which is about 750 miles per hour.

a. If you see a flash of lightning, then hear the thunder 4 seconds later, how far away is the lightning?

b. Now let's generalize that result. Suppose it takes n seconds to hear the thunder after a flash of lightning. How far away is the lightning, in terms of n?

60. Sound travels about 750 miles per hour. If you stand in a parking lot near a building and sound a horn, you will hear an echo.

a. Suppose it takes about $\frac{1}{2}$ a second to hear the echo. How far away is the building[8]?

b. Now let's generalize that result. Suppose it takes n seconds to hear the echo. How far away is the building, in terms of n?

61. It takes an air pump 5 minutes to fill a twin sized air mattress (39 by 8.75 by 75 inches). How long will it take to fill a queen sized mattress (60 by 8.75 by 80 inches)?

62. It takes your garden hose 20 seconds to fill your 2-gallon watering can. How long will it take to fill

a. An inflatable pool measuring 3 feet wide, 8 feet long, and 1 foot deep.[9]

b. A circular inflatable pool 13 feet in diameter and 3 feet deep.[10]

63. You want to put a 2" thick layer of topsoil for a new 20'x30' garden. The dirt store sells by the cubic yards. How many cubic yards will you need to order?

64. A box of Jell-O costs $0.50, and makes 2 cups. How much would it cost to fill a swimming pool 4 feet deep, 8 feet wide, and 12 feet long with Jell-O? (1 cubic foot is about 7.5 gallons)

65. You read online that a 15 ft by 20 ft brick patio would cost about $2,275 to have professionally installed. Estimate the cost of having a 18 by 22 ft brick patio installed.

66. I was at the store, and saw two sizes of avocados being sold. The regular size sold for $0.88 each, while the jumbo ones sold for $1.68 each. Which is the better deal?

67. The grocery store has bulk pecans on sale, which is great since you're planning on making 10 pecan pies for a wedding. Your recipe calls for 1¾ cups pecans per pie. However, in the bulk section there's only a scale available, not a measuring cup. You run over to the baking aisle and find a bag of pecans, and look at the nutrition label to gather some info. How many pounds of pecans should you buy?

Nutrition Facts

Serving Size: 1 cup, halves (99 g)
Servings per Container: about 2

Amount Per Serving	
Calories 684	Calories from Fat 596
	% Daily Value*
Total Fat 71g	110%
Saturated Fat 6g	31%
Trans Fat	
Cholesterol 0mg	0%

68. Soda is often sold in 20 ounce bottles. The nutrition label for one of these bottles is shown to the right. A packet of sugar (the kind they have at restaurants for your coffee or tea) typically contain 4 grams of sugar in the U.S. Drinking a 20 oz soda is equivalent to eating how many packets of sugar?[11]

Nutrition Facts	
Serving Size: 8 fl oz (240 mL)	
Servings Per Container: about 2.5	
Amount Per Serving	
Calories 110	
	% Daily Value*
Total Fat 0g	0%
Sodium 70mg	3%
Total Carbohydrate 31g	10%
Sugars 30g	
Protein 0g	

For the next set of questions, *first* identify the information you need to answer the question, and *then* turn to the end of the section to find that information. The details may be imprecise; answer the question the best you can with the provided information. Be sure to justify your decision.

69. You're planning on making 6 meatloafs for a party. You go to the store to buy breadcrumbs, and see they are sold by the canister. How many canisters do you need to buy?

70. Your friend wants to cover their car in bottle caps, like in this picture.[12] How many bottle caps are you going to need?

71. You need to buy some chicken for dinner tonight. You found an ad showing that the store across town has it on sale for $2.99 a pound, which is cheaper than your usual neighborhood store, which sells it for $3.79 a pound. Is it worth the extra drive?

72. I have an old gas furnace, and am considering replacing it with a new, high efficiency model. Is upgrading worth it?

73. Janine is considering buying a water filter and a reusable water bottle rather than buying bottled water. Will doing so save her money?

74. Marcus is considering going car-free to save money and be more environmentally friendly. Is this financially a good decision?

For the next set of problems, research or make educated estimates for any unknown quantities needed to answer the question.

75. You want to travel from Tacoma, WA to Chico, CA for a wedding. Compare the costs and time involved with driving, flying, and taking a train. Assume that if you fly or take the train you'll need to rent a car while you're there. Which option is best?

76. You want to paint the walls of a 6ft by 9ft storage room that has one door and one window. You want to put on two coats of paint. How many gallons and/or quarts of paint should you buy to paint the room as cheaply as possible?

77. A restaurant in New York tiled their floor with pennies[13]. Just for the materials, is this more expensive than using a more traditional material like ceramic tiles? If each penny has to be laid by hand, estimate how long it would take to lay the pennies for a 12ft by 10ft room. Considering material and labor costs, are pennies a cost-effective replacement for ceramic tiles?

78. You are considering taking up part of your back yard and turning it into a vegetable garden, to grow broccoli, tomatoes, and zucchini. Will doing so save you money, or cost you more than buying vegetables from the store?

79. Barry is trying to decide whether to keep his 1993 Honda Civic with 140,000 miles, or trade it in for a used 2008 Honda Civic. Consider gas, maintenance, and insurance costs in helping him make a decision.

80. Some people claim it costs more to eat vegetarian, while some claim it costs less. Examine your own grocery habits, and compare your current costs to the costs of switching your diet (from omnivore to vegetarian or vice versa as appropriate). Which

diet is more cost effective based on your eating habits?

Info for the breadcrumbs question

How much breadcrumbs does the recipe call for?

It calls for 1½ cups of breadcrumbs.

How many meatloafs does the recipe make?

It makes 1 meatloaf.

How many servings does that recipe make?

It says it serves 8.

How big is the canister?

It is cylindrical, 3.5 inches across and 7 inches tall.

What is the net weight of the contents of 1 canister?

15 ounces.

How much does a cup of breadcrumbs weigh?

I’m not sure, but maybe something from the nutritional label will help.

Nutrition Facts

Serving Size: 1/3 cup (30g)
Servings per Container: about 14

Amount Per Serving	
Calories 110	Calories from Fat 15
	% Daily Value*
Total Fat 1.5g	2%

How much does a canister cost?

$2.39

Info for bottle cap car

What kind of car is that?

A 1993 Honda Accord.

How big is that car / what are the dimensions? Here is some details from MSN autos:

Weight: 2800lb Length: 185.2 in Width: 67.1 in Height: 55.2 in

How much of the car was covered with caps?

Everything but the windows and the underside.

How big is a bottle cap?

Caps are 1 inch in diameter.

Info for chicken problem

How much chicken will you be buying?

Four pounds

How far are the two stores?

My neighborhood store is 2.2 miles away, and takes about 7 minutes. The store across town is 8.9 miles away, and takes about 25 minutes.

What kind of mileage does your car get?

It averages about 24 miles per gallon in the city.

How many gallons does your car hold?

About 14 gallons

How much is gas?

About $3.69/gallon right now.

Info for furnace problem

How efficient is the current furnace?

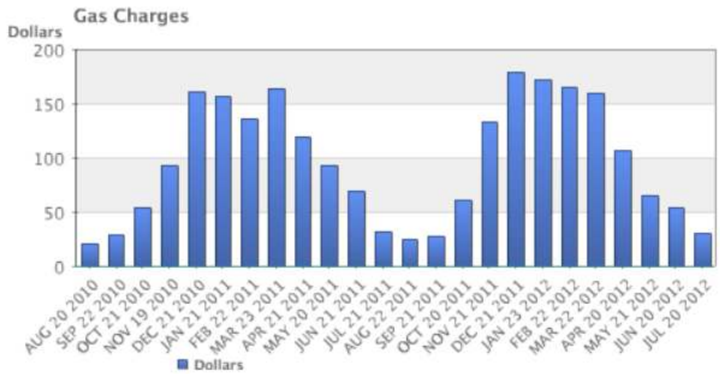
It is a 60% efficient furnace.

How efficient is the new furnace?

It is 94% efficient.

What is your gas bill?

Here is the history for 2 years:

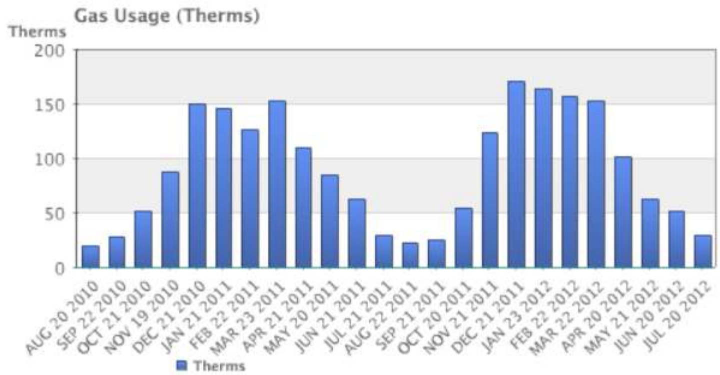


How much do you pay for gas?

There is $10.34 base charge, plus $0.39097 per Therm for a delivery charge, and $0.65195 per Therm for cost of gas.

How much gas do you use?

Here is the history for 2 years:



How much does the new furnace cost?

It will cost $7,450.

How long do you plan to live in the house?

Probably at least 15 years.

Info for water filter problem

How much water does Janine drink in a day?

She normally drinks 3 bottles a day, each 16.9 ounces.

How much does a bottle of water cost?

She buys 24-packs of 16.9 ounce bottles for $3.99.

How much does a reusable water bottle cost?

About $10.

How long does a reusable water bottle last?

Basically forever (or until you lose it).

How much does a water filter cost? How much water will they filter?

- A faucet-mounted filter costs about $28. Refill filters cost about $33 for a 3-pack. The box says each filter will filter up to 100 gallons (378 liters)
- A water filter pitcher costs about $22. Refill filters cost about $20 for a 4-pack. The box says each filter lasts for 40 gallons or 2 months
- An under-sink filter costs $130. Refill filters cost about $60 each. The filter lasts for 500 gallons.

Info for car-free problem

Where does Marcus currently drive? He:

- Drives to work 5 days a week, located 4 miles from his house.
- Drives to the store twice a week, located 7 miles from his house.
- Drives to other locations on average 5 days a week, with locations ranging from 1 mile to 20 miles.
- Drives to his parent's house 80 miles away once a month.

How will he get to these locations without a car?

- For work, he can walk when it's sunny and he gets up early enough. Otherwise he can take a bus, which takes about 20 minutes
- For the store, he can take a bus, which takes about 35 minutes.
- Some of the other locations he can bus to. Sometimes he'll be able to get a friend to pick him up. A few locations he is able to walk to. A couple locations are hard to get to by bus, but there is a ZipCar (short term car rental) location within a few blocks.
- He'll need to get a ZipCar to visit his parents.

How much does gas cost?

About $3.69/gallon.

How much does he pay for insurance and maintenance?

- He pays $95/month for insurance.
- He pays $30 every 3 months for an oil change, and has averaged about $300/year for other maintenance costs.

How much is he paying for the car?

- He's paying $220/month on his car loan right now, and has 3 years left on the loan.
- If he sold the car, he'd be able to make enough to pay off the loan.
- If he keeps the car, he's planning on trading the car in for a newer model in a couple years.

What mileage does his car get?

About 26 miles per gallon on average.

How much does a bus ride cost?

$2.50 per trip, or $90 for an unlimited monthly pass.

How much does a ZipCar rental cost?

- The "occasional driving plan": $25 application fee and $60 annual fee, with no monthly commitment. Monday-Thursday the cost is $8/hour, or $72 per day. Friday-Sunday the cost is $8/hour or $78/day. Gas, insurance, and 180 miles are included in the cost. Additional miles are $0.45/mile.
- The "extra value plan": Same as above, but with a $50 monthly commitment, getting you a 10% discount on the usage costs.

[1] http://www.cnn.com/2010/SHOWBIZ/Musi...les/index.html

[2] http://www.cnn.com/2012/06/27/politi...are/index.html

[3] http://www.politico.com/news/stories/0712/78134.html

[4] http://factcheck.org/2012/07/twistin...th-care-taxes/

[5] www.northjersey.com/news/opin...tml?c=y&page=2

[6] http://www.the-numbers.com/movies/records/budgets.php

[7] vimeo.com/42501010

[8] vimeo.com/40377128

[9] http://www.youtube.com/watch?v=DIkwefReHZc

[10] http://www.youtube.com/watch?v=p9SABH7Yg9M

[11] http://www.youtube.com/watch?v=62JMfv0tf3Q

[12] Photo credit: http://www.flickr.com/photos/swayze/, CC-BY

[13] http://www.notcot.com/archives/2009/...-of-pennie.php

1.7: Extension - Taxes

Governments collect taxes to pay for the services they provide. In the United States, federal income taxes help fund the military, the environmental protection agency, and thousands of other programs. Property taxes help fund schools. Gasoline taxes help pay for road improvements. While very few people enjoy paying taxes, they are necessary to pay for the services we all depend upon.

Taxes can be computed in a variety of ways, but are typically computed as a percentage of a sale, of one's income, or of one's assets.

Example 1

The sales tax rate in a city is 9.3%. How much sales tax will you pay on a $140 purchase?

Solution

The sales tax will be 9.3% of $140. To compute this, we multiply $140 by the percent written as a decimal: $\$140(0.093)=\$13.02$

When taxes are not given as a fixed percentage rate, sometimes it is necessary to calculate the **effective rate.**

Effective rate

The effective tax rate is the equivalent percent rate of the tax paid out of the dollar amount the tax is based on.

Example 2

Joan paid $3,200 in property taxes on her house valued at $215,000 last year. What is the effective tax rate?

Solution

We can compute the equivalent percentage: $3200/215000=0.01488$or about 1.49% effective rate.

Taxes are often referred to as progressive, regressive, or flat.

Tax categories

A **flat tax**, or proportional tax, charges a constant percentage rate.

A **progressive tax** increases the percent rate as the base amount increases.

A **regressive tax** decreases the percent rate as the base amount increases.

Example 3

The United States federal income tax on earned wages is an example of a progressive tax. People with a higher wage income pay a higher percent tax on their income.

Solution

For a single person in 2011, adjusted gross income (income after deductions) under $8,500 was taxed at 10%. Income over $8,500 but under $34,500 was taxed at 15%.

A person earning $10,000 would pay 10% on the portion of their income under $8,500, and 15% on the income over $8,500, so they'd pay:

$$\begin{array}{ll} 8500(0.10)=850 & 10\% \text{ of } 8500 \\ 1500(0.15)=225 & 15\% \text{ of the remaining } \$1500 \text{ of income} \\ \text{Total tax: } =\$1075 & \end{array}$$

The effective tax rate paid is $1075/10000=10.75\%$

A person earning \$30,000 would also pay 10% on the portion of their income under \$8,500, and 15% on the income over \$8,500, so they'd pay:

$$\begin{array}{ll} 8500(0.10)=850 & 10\% \text{ of } 8500 \\ 21500(0.15)=3225 & 15\% \text{ of the remaining } \$21500 \text{ of income} \\ \text{Total tax: } =\$4075 & \end{array}$$

The effective tax rate paid is $4075/30000 = 13.58\%$

Notice that the effective rate has increased with income, showing this is a progressive tax.

✔ Example 4

A gasoline tax is a flat tax when considered in terms of consumption, a tax of, say, \$0.30 per gallon is proportional to the amount of gasoline purchased. Someone buying 10 gallons of gas at \$4 a gallon would pay \$3 in tax, which is $\$3/\$40 = 7.5\%$ Someone buying 30 gallons of gas at \$4 a gallon would pay \$9 in tax, which is $\$9/\$120 = 7.5\%$ the same effective rate.

Solution

However, in terms of income, a gasoline tax is often considered a regressive tax. It is likely that someone earning \$30,000 a year and someone earning \$60,000 a year will drive about the same amount. If both pay \$60 in gasoline taxes over a year, the person earning \$30,000 has paid 0.2% of their income, while the person earning \$60,000 has paid 0.1% of their income in gas taxes.

? Try it Now 1

A sales tax is a fixed percentage tax on a person's purchases. Is this a flat, progressive, or regressive tax?

Answer

While sales tax is a flat percentage rate, it is often considered a regressive tax for the same reasons as the gasoline tax.

1.8: Income Taxation

Many people have proposed various revisions to the income tax collection in the United States. Some, for example, have claimed that a flat tax would be fairer. Others call for revisions to how different types of income are taxed, since currently investment income is taxed at a different rate than wage income.

The following two projects will allow you to explore some of these ideas and draw your own conclusions.

Project 1: Flat tax, Modified Flat Tax, and Progressive Tax.

Imagine the country is made up of 100 households. The federal government needs to collect $800,000 in income taxes to be able to function. The population consists of 6 groups:

Group A: 20 households that earn $12,000 each

Group B: 20 households that earn $29,000 each

Group C: 20 households that earn $50,000 each

Group D: 20 households that earn $79,000 each

Group E: 15 households that earn $129,000 each

Group F: 5 households that earn $295,000 each

This scenario is roughly proportional to the actual United States population and tax needs. We are going to determine new income tax rates.

The first proposal we'll consider is a flat tax – one where every income group is taxed at the same percentage tax rate.

1) Determine the total income for the population (all 100 people together)

2) Determine what flat tax rate would be necessary to collect enough money.

The second proposal we'll consider is a modified flat-tax plan, where everyone only pays taxes on any income over $20,000. So, everyone in group A will pay no taxes. Everyone in group B will pay taxes only on $9,000.

3) Determine the total *taxable* income for the whole population

4) Determine what flat tax rate would be necessary to collect enough money in this modified system

5) Complete this table for both the plans

		Flat Tax Plan		Modified Flat Tax Plan	
Group	**Income per household**	**Income tax per household**	**Income after taxes**	**Income tax per household**	**Income after taxes**
A	$12,000				
B	$29,000				
C	$50,000				
D	$79,000				
E	$129,000				
F	$295,000				

The third proposal we'll consider is a progressive tax, where lower income groups are taxed at a lower percent rate, and higher income groups are taxed at a higher percent rate. For simplicity, we're going to assume that a household is taxed at the same rate on *all* their income.

6) Set progressive tax rates for each income group to bring in enough money. There is no one right answer here – just make sure you bring in enough money!

Group	Income per household	Tax rate (%)	Income tax per household	Total tax collected for all households	Income after taxes per household
A	$12,000				
B	$29,000				
C	$50,000				
D	$79,000				
E	$129,000				
F	$295,000				
				This better total to $800,000	

7) Discretionary income is the income people have left over after paying for necessities like rent, food, transportation, etc. The cost of basic expenses does increase with income, since housing and car costs are higher, however usually not proportionally. For each income group, estimate their essential expenses, and calculate their discretionary income. Then compute the effective tax rate for each plan relative to discretionary income rather than income.

Group	Income per household	Discretionary Income (estimated)	Effective rate, flat	Effective rate, modified	Effective rate, progressive
A	$12,000				
B	$29,000				
C	$50,000				
D	$79,000				
E	$129,000				
F	$295,000				

8) Which plan seems the most fair to you? Which plan seems the least fair to you? Why?

Project 2: Calculating Taxes.

Visit www.irs.gov, and download the most recent version of forms 1040, and schedules A, B, C, and D.

Scenario 1: Calculate the taxes for someone who earned $60,000 in standard wage income (W-2 income), has no dependents, and takes the standard deduction.

Scenario 2: Calculate the taxes for someone who earned $20,000 in standard wage income, $40,000 in qualified dividends, has no dependents, and takes the standard deduction. (Qualified dividends are earnings on certain investments such as stocks.)

Scenario 3: Calculate the taxes for someone who earned $60,000 in small business income, has no dependents, and takes the standard deduction.

Based on these three scenarios, what are your impressions of how the income tax system treats these different forms of income (wage, dividends, and business income)?

Scenario 4: To get a more realistic sense for calculating taxes, you'll need to consider itemized deductions. Calculate the income taxes for someone with the income and expenses listed below.

Married with 2 children, filing jointly

Wage income: $50,000 combined

Paid sales tax in Washington State

Property taxes paid: $3200

Home mortgage interest paid: $4800

Charitable gifts: $1200

CHAPTER OVERVIEW

2: Set Theory

It is natural for us to classify items into groups, or sets, and consider how those sets overlap with each other. We can use these sets understand relationships between groups, and to analyze survey data.

via Wikipedia).

2.1: Language of Sets

An art collector might own a collection of paintings, while a music lover might keep a collection of CDs. Any collection of items can form a **set**.

Set

A **set** is a collection of distinct objects, called **elements** or **members** of the set

A set can be defined by listing the elements of the set, or by describing the contents in set-builder notation, enclosed in curly brackets.

Example 1

Some examples of sets defined by **listing** the elements of the set:

a. A = {1, 3, 9, 12}
b. B = {red, orange, yellow, green, blue, indigo, purple}

Some examples of sets defined by describing the contents in **set-builder notation**:

a. C = {x : x is an even number}
b. D = {y : y is a book written about travel to Chile}

A set simply specifies the contents; order is not important. The set represented by {1, 2, 3} is the same as the set {3, 1, 2}.

Set Notation

Commonly, we will use a variable to represent a set, to make it easier to refer to that set later.

The symbol $\in$ means "is an element of".

A set that contains no elements, $\{\}$, is called the **empty set** or **null set** and is notated $\emptyset$

Example 2

Let $A=\{1,2,3,4\}$

To notate that 2 is element of the set, we'd write $2 \in A$

Well-defined

A set is **well-defined** if we are able to tell whether any particular object is an element of the set.

Example 3

Which sets are well-defined and which sets are not well-defined?

a. A = {b : b is a type of tree}
b. B = {g : g is a tasty food}
c. C = {z : z is a restaurant in San Francisco}

Solution

The sets A and C are well-defined because we know exactly what types of trees there are and restaurants in San Francisco. The set B is not well-defined because there are different ideas of what a tasty food is.

Universal Set

The **universal set** is the set of all elements under consideration in a given discussion denoted by the letter U.

Example: U = {k : k is a student at Las Positas College}

Often times we are interested in the number of items in a set. This is called the cardinality of the set.

Cardinal Number

The number of elements in a set is the **cardinal number** of that set.

The cardinal number of the set A is often notated as $n(A)$

Example 4

What is the cardinal number of $\emptyset$?

Solution

Since this is the empty set, $n(\emptyset)$=0

Example 5

What is the cardinal number of P = {h : h is the English name for the months of the year}?

Solution

The cardinal number, $n(P)$=12, since there are 12 months in the year

---

2.2: Comparing Sets

Equal and Equivalent

Two sets A and B are **equal** if they have exactly the same elements. We write A = B.

Two sets A and B are **equivalent** if $n(A) = n(B)$. Another way of saying this is that two sets are equivalent if they have the same number of elements.

Example 1

Determine if the following pairs of sets are equal, equivalent, or both

a. {1, 3, 5, 7, 9} and {7, 5, 1, 3, 9}
b. ∅ and {x : x is a living human born before 1600}
c. {Facebook, Instagram, Tik Tok, Snapchat} and {baseball, hockey, football, basketball}

Solution

a. These two sets are equal because they have the same exact elements. It does not matter that they are rearranged. They are also equivalent because they both have 5 elements.
b. These two sets are equal because there are no humans alive that were born before 1600. They are also equivalent because they both have 0 elements.
c. These two sets are not equal because one set is social media platforms and the other set is sports. They are, however, equivalent because they both have 4 elements.

Sometimes a collection might not contain all the elements of a set. For example, Chris owns three Madonna albums. While Chris's collection is a set, we can also say it is a **subset** of the larger set of all Madonna albums.

Subset

A **subset** of a set A is another set that contains only elements from the set A, but may not contain all the elements of A.

If B is a subset of A, we write $B \subseteq A$

A **proper subset** is a subset that is not identical to the original set – it contains fewer elements.

If B is a proper subset of A, we write $B \subset A$

Example 2

Consider these three sets

$A =$ the set of all even numbers $\quad B = \{2, 4, 6\} \quad C = \{2, 3, 4, 6\}$

Here $B \subset A$ since every element of B is also an even number, so is an element of A.

More formally, we could say $B \subset A$ since if $x \in B$, then $x \in A$

It is also true that $B \subset C$.

C is not a subset of A, since C contains an element, 3 , that is not contained in A

Example 3

Suppose a set contains the plays "Much Ado About Nothing", "MacBeth", and "A Midsummer's Night Dream". What is a larger set this might be a subset of?

Solution

There are many possible answers here. One would be the set of plays by Shakespeare. This is also a subset of the set of all plays ever written. It is also a subset of all British literature.

Try it Now 1

The set $A = \{1, 3, 5\}$. What is a larger set this might be a subset of?

Answer

There are several answers: The set of all odd numbers less than 10. The set of all odd numbers. The set of all integers. The set of all real numbers.

One way to build intuition about subsets is to try listing all the different subsets of a particular set. Let's look at some examples of small sets and identify all of their subsets.

Example 4

List all of the subsets for the following sets:

a. The empty set ∅

b. {a}

c. {m, n}

d. {x, y, z}

Solution

a. Since the empty set has zero elements, the only subset of the empty set ∅ is the empty set itself.

b. The set {a} has two subsets: the set {a} itself as well as the empty set ∅

c. The set {m, n} has four subsets: the empty set ∅, {m}, {n} and {m, n}

d. The set {x, y, z} has eight subsets: the empty set ∅, {x}, {y}, {z}, {x, y}, {x, z}, {y, z}, and {x, y, z}

There are several things to observe here. First notice that every set has itself as a subset. Also, the empty set is a subset of every possible set. Lastly, did you see what happened as the sets increased in size? For each new element of the set - 0, 1, 2, 3 - the number of subsets doubles - 1, 2, 4, 8. This pattern continues for sets of any size, so we can come up with a formula to predict the number of subsets for a given set.

Definition: Number of Subsets of a Set

A set with **k** elements has $\mathbf{2^k}$ different subsets.

2.3: Set Operations

Commonly sets interact. For example, you and a new roommate decide to have a house party, and you both invite your circle of friends. At this party, two sets are being combined, though it might turn out that there are some friends that were in both sets.

Union, Intersection, and Complement

The **union** of two sets contains all the elements contained in either set (or both sets).

The union is notated $A \cup B$

More formally, $x \in A \cup B$ if $x \in A$ or $x \in B$ (or both)

The **intersection** of two sets contains only the elements that are in both sets.

The intersection is notated $A \cap B$

More formally, $x \in A \cap B$ if $x \in A$ and $x \in B$

If *A* is a subset of the universal set U, then the **complement** of *A* is the set of elements that are elements of U, but are *not* in the set *A*.

The complement is notated A', or A^c, or sometimes $\sim A$.

More formally, $x \in A'$ if $x \in U$ and $x \notin A$

The **difference** of sets *A* and *B* is the set of elements in *A* that are *not* in *B*.

The difference is notated $A - B$

More formally, $x \in A - B$ if $x \in A$ and $x \notin B$

Example 1

Consider the sets:

$U = \{\text{red, orange, yellow, green, blue, indigo, purple}\} \quad A = \{\text{ red, green, blue }\} \quad B = \{\text{ red, yellow, orange }\}$
$C = \{\text{ red, orange, yellow, green, blue, purple }\}$

a. Find $A \cup B$
b. Find $A \cap B$
c. Find $A' \cap C$
d. Find $A - B$

Solution

a) The union contains all the elements in either set: $A \cup B = \{\text{ red, green, blue, yellow, orange }\}$

Notice we only list red once.

b) The intersection contains all the elements in both sets: $A \cap B = \{\text{ red }\}$

c) Here we're looking for all the elements that are not in set A and are also in C.

$A' \cap C = \{\text{ orange, yellow, purple }\}$

d) Here we want to find all the elements that are in the set A that are not in B.

$A - B = \{\text{ green, blue}\}$

Try it Now 1

Using the sets from the previous example, find $A \cup C$ and $B' \cap A$

Answer

$A \cup C = \{\text{ red, orange, yellow, green, blue purple }\}$

$B' \cap A = \{ \text{ green, blue } \}$

Example 2

Suppose the universal set is $U =$ all whole numbers from 1 to 9. If $A = \{1, 2, 4\}$, then

$A' = \{3, 5, 6, 7, 8, 9\}$

As we saw earlier with the expression $A^c \cap C$, set operations can be grouped together. Grouping symbols can be used like they are with arithmetic - to force an order of operations.

Example 3

Suppose

$H = \{ \text{ cat, dog, rabbit, mouse } \}, F = \{ \text{ dog, cow, duck, pig, rabbit } \} \quad W = \{ \text{ duck, rabbit, deer, frog, mouse } \}$

a. Find $(H \cap F) \cup W$
b. Find $H \cap (F \cup W)$
c. Find $(H \cap F)' \cap W$

Solution

a) We start with the intersection: $H \cap F = \{ \text{ dog, rabbit } \}$

Now we union that result with $W : (H \cap F) \cup W = \{\text{dog, duck, rabbit, deer, frog, mouse } \}$

b) We start with the union: $F \cup W = \{\text{dog, cow, rabbit, duck, pig, deer, frog, mouse } \}$

Now we intersect that result with $H : H \cap (F \cup W) = \{ \text{ dog, rabbit, mouse } \}$

c) We start with the intersection: $H \cap F = \{\text{dog, rabbit } \}$

Now we want to find the elements of W that are not in H $\cap F$

$(H \cap F)' \cap W = \{ \text{ duck, deer, frog, mouse } \}$

To visualize the interaction of sets, John Venn in 1880 thought to use overlapping circles, building on a similar idea used by Leonhard Euler in the 18$^{\text{th}}$ century. These illustrations now called **Venn Diagrams**.

Venn Diagram

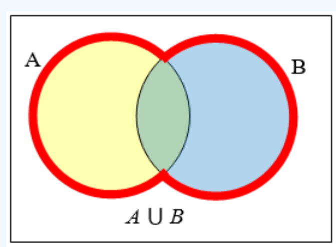
A Venn diagram represents each set by a circle, usually drawn inside of a containing box representing the universal set. Overlapping areas indicate elements common to both sets.

Basic Venn diagrams can illustrate the interaction of two or three sets.

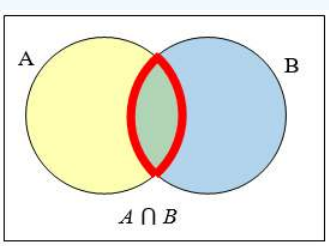
Example 4

Create Venn diagrams to illustrate $A \cup B$, $A \cap B$, and $A' \cap B$

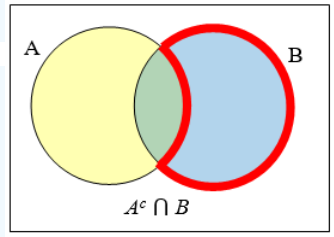
$A \cup B$ contains all elements in either set.



$A \cup B$ contains all elements in either set.



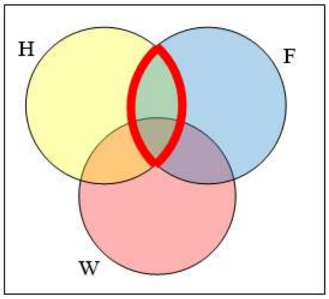
$A \cap B$ contains only those elements in both sets - in the overlap of the circles.



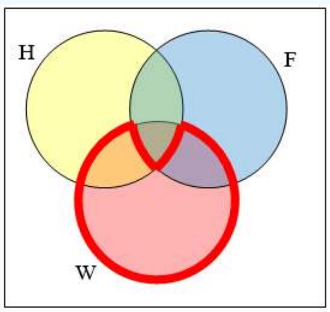
Example 5

Use a Venn diagram to illustrate $(H \cap F)' \cap W$

We'll start by identifying everything in the set $\mathrm{H} \cap F$



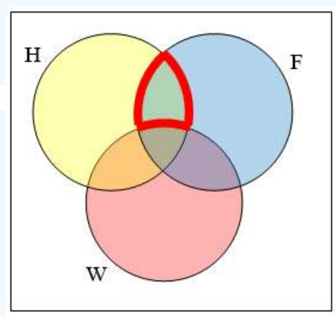
Now, $(H \cap F)' \cap W$ will contain everything not in the set identified above that is also in set W



Example 6

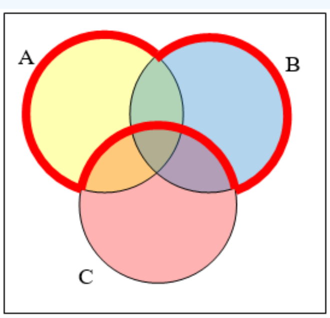
Create an expression to represent the outlined part of the Venn diagram shown.

The elements in the outlined set are in sets H and F, but are not in set W. So we could represent this set as $H \cap F \cap W$



Try it Now 2

Create an expression to represent the outlined portion of the Venn diagram shown



Answer

$A \cup B \cap C'$

DeMorgan's Laws

If A and B are sets, then $(A \cup B)' = A' \cap B'$ and $(A \cap B)' = A' \cup B'$

Example 7

Let $A=\{1,2,3,4,5\}$ and $B=\{2,4,6,8,10\}$, with universal set $U=\{1,2,3,4,5,6,7,8,9,10\}$ Verify one of the above laws by showing that $(A \cup B)' = A' \cap B'$

Solution

$A \cup B=\{1,2,3,4,5,6,8,10\}$ and so $A \cup B' =\{7,9\}$.

$A'=\{6,7,8,9,10\}$ and $B'=\{1,3,5,7,9\}$, so $A' \cap B'=\{7,9\}$. This confirms that both sides are equal and $(A \cup B)' = A' \cap B'$.

2.4: Survey Problems

Sometimes we may be interested in the cardinality of the union or intersection of sets, but not know the actual elements of each set. This is common in surveying.

Example 1

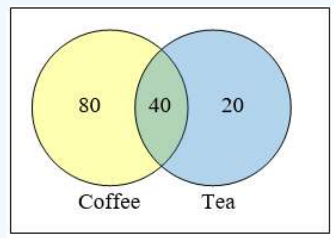
A survey asks 200 people "What beverage do you drink in the morning", and offers choices:

- Tea only
- Coffee only
- Both coffee and tea

Suppose 20 report tea only, 80 report coffee only, 40 report both. How many people drink tea in the morning? How many people drink neither tea or coffee?

Solution

This question can most easily be answered by creating a Venn diagram. We can see that we can find the people who drink tea by adding those who drink only tea to those who drink both: 60 people.



We can also see that those who drink neither are those not contained in the any of the three other groupings, so we can count those by subtracting from the cardinality of the universal set, 200.

$200-20-80-40=60$ people who drink neither.

Example 2

A survey asks: Which online services have you used in the last month:

- Twitter
- Facebook
- Have used both

The results show 40% of those surveyed have used Twitter, 70% have used Facebook, and 20% have used both. How many people have used neither Twitter or Facebook?

Solution

Let T be the set of all people who have used Twitter, and F be the set of all people who have used Facebook. Notice that while the cardinality of F is 70% and the cardinality of T is 40%, the cardinality of $F \cup T$ is not simply $70\% + 40\%$, since that would count those who use both services twice. To find the cardinality of $F \cup T$, we can add the cardinality of F and the cardinality of T, then subtract those in intersection that we've counted twice. In symbols,

$\mathrm{n}(F \cup T)=\mathrm{n}(F)+\mathrm{n}(T)-\mathrm{n}(F \cap T)$

$\mathrm{n}(F \cup T)=70\%+40\%-20\%=90\%$

Now, to find how many people have not used either service, we're looking for the cardinality of $(F \cup T)'$. since the universal set contains 100% of people and the cardinality of $F \cup T=90\%$, the cardinality of $(F \cup T)'$ must be the other 10%

The previous example illustrated two important properties

Cardinality Properties

$\mathrm{n}(A \cup B)=\mathrm{n}(A)+\mathrm{n}(B)-\mathrm{n}(A \cap B)$

$n\left(A'\right)=n(U)-n(A)$

Example 3

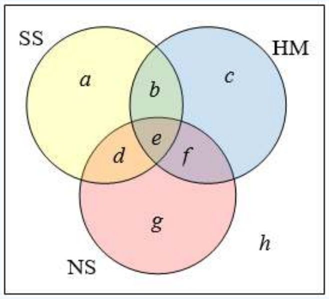
Fifty students were surveyed, and asked if they were taking a social science (SS), humanities (HM) or a natural science (NS) course the next quarter.

21 were taking a SS course	26 were taking a HM course
19 were taking a NS course	9 were taking SS and HM
7 were taking SS and NS	10 were taking HM and NS
3 were taking all three	7 were taking none

How many students are only taking a SS course?

Solution

It might help to look at a Venn diagram.



From the given data, we know that there are 3 students in region e and 7 students in region h

since 7 students were taking a SS and NS course, we know that $n(d)+n(e)=7$. since we know there are 3 students in region e , there must be $7-3=4$ students in region d

Similarly, since there are 10 students taking HM and NS, which includes regions e and f, there must be $10-3=7$ students in region f

Since 9 students were taking SS and HM, there must be $9-3=6$ students in region b

Now, we know that 21 students were taking a SS course. This includes students from regions a, b, d, and e. since we know the number of students in all but region a, we can determine that $21-6-4-3=8$ students are in region a

8 students are taking only a SS course.

Try it Now 1

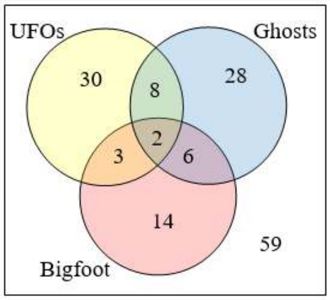
One hundred fifty people were surveyed and asked if they believed in UFOs, ghosts, and Bigfoot.

43 believed in UFOs	44 believed in ghosts
25 believed in Bigfoot	10 believed in UFOs and ghosts
8 believed in ghosts and Bigfoot	5 believed in UFOs and Bigfoot
2 believed in all three	

How many people surveyed believed in none of these things?

Answer

Starting with the intersection of all three circles, we work our way out. since 10 people believe in UFOs and Ghosts, and 2 believe in all three, that leaves 8 that believe in only UFOs and Ghosts. We work our way out, filling in all the regions. Once we have, we can add up all those regions, getting 91 people in the union of all three sets. This leaves $150-91=59$ who believe in none.



CHAPTER OVERVIEW

3: Logic

Thumbnail: pixabay.com/illustrations/pu...ution-1721464/

3.1: Statements, Connectives, and Quantifiers

Statements

A **statement** in logic is a declarative sentence that is either true or false. We represent statements by lowercase letters such as *p*, *q*, or *r*.

Examples of statements:

1. Guillermo has the flu.
2. Either George Washington or Abraham Lincoln was the first president of the United States
3. If Paul eats pizza, then Paul will drink soda.

Examples of Non-statements:

1. Go clean your room. (This is not declaring a truth or false)
2. Why are you here? (This is a question)
3. This sentence is false. (This is a paradox. It cannot be either true or false. If we assume that this sentence is true, then we must conclude that it is false. If we assume that it is false, then we must conclude that it is true.)

It is possible to form new statements from existing statements by connecting the statements with words such as “and” and “or” or by negating the statement.

Compound Statements and Connectives

A **connective** on a statement is a word or combination of words that combines one or more statements to make a new mathematical statement.

A **compound statement** is a statement that contains one or more connectives. Because these connectives are used so frequently in logic, we give them names and use special symbols to represent them.

Connectives

A **negation** expresses the word "not" and uses the symbol $\sim$: not p is notated $\sim p$

A **conjunction** expresses the word "and" between two statements and uses the symbol $\wedge$: p and q is notated $p \wedge q$

A **disjunction** expresses the word "or" between two statements and uses the symbol $\vee$: p or q is notated $p \vee q$

A **conditional** statement is of the form "If ... then ..." and uses the symbol $\rightarrow$: If p, then q is notated $p \rightarrow q$

A **biconditional** statement expresses "if and only if" between two statements and uses the symbol $\leftrightarrow$: p if and only if q is notated $p \leftrightarrow q$

You can remember the first two symbols by relating them to the shapes for the union and intersection. $A \wedge B$ would be the elements that exist in both sets, in $A \cap B$. Likewise, $A \vee B$ would be the elements that exist in either set, in $A \cup B$. When we are working with sets, we use the rounded version of the symbols; when we are working with statements, we use the pointy version.

Example 1

Translate each statement into symbolic notation. Let P represent "I like Pepsi" and let C represent "I like Coke."

a. I like Pepsi or I like Coke.
b. I like Pepsi and I like Coke.
c. I do not like Pepsi.
d. It is not the case that I like Pepsi or Coke.
e. I like Pepsi and I do not like Coke.

Solution

a. $P \vee C$
b. $P \wedge C$
c. $\sim P$
d. $\sim (P \vee C)$
e. $P \wedge \sim C$

As you can see, we can use parentheses to organize more complicated statements.

Try it Now 1

Translate "We have carrots or we will not make soup" into symbols. Let C represent "we have carrots" and let S represent "we will make soup".

Answer

$C \vee \sim S$

Quantifiers

A **universal quantifier** states that an entire set of things share a characteristic. (all, every, or none)

An **existential quantifier** states that a set contains at least one element. (some, many, or at least one)

Something interesting happens when we **negate** – or state the opposite of – a quantified statement.

Example 2

Suppose your friend says "Everybody cheats on their taxes." What is the minimum amount of evidence you would need to prove your friend wrong?

To show that it is not true that everybody cheats on their taxes, all you need is one person who does not cheat on their taxes. It would be perfectly fine to produce more people who do not cheat, but one counterexample is all you need.

It is important to note that you do not need to show that absolutely nobody cheats on their taxes.

Example 3

Suppose your friend says "One of these six cartons of milk is leaking." What is the minimum amount of evidence you would need to prove your friend wrong?

Solution

In this case, you would need to check all six cartons and show that none of them is leaking. You cannot disprove your friend's statement by checking only one of the cartons.

When we negate a statement with a universal quantifier, we get a statement with an existential quantifier, and vice-versa.

Negating a Quantified Statement

The negation of "all A are B" is "at least one A is not B".

The negation of "no A are B" is "at least one A is B".

The negation of "at least one A is B" is "no A are B".

The negation of "at least one A is not B" is "all A are B".

Example 4

"Somebody brought a flashlight." Write the negation of this statement.

The negation is "Nobody brought a flashlight."

Example 5

"There are no prime numbers that are even." Write the negation of this statement.

The negation is "At least one prime number is even."

Try it Now 2

Write the negation of "All Icelandic children learn English in school."

Answer

At least one Icelandic child did not learn English in school.

3.2: Truth Tables- Conjunction (and), Disjunction (or), Negation (not)

Because complex statements can get tricky to think about, we can create a **truth table** to keep track of what truth values for the simple statements make the complex statement true and false.

Truth Table

A table showing what the resulting truth value of a complex statement is for all the possible truth values for the simple statements.

Example 1

Suppose you're picking out a new couch, and your significant other says "get a sectional *or* something with a chaise".

This is a complex statement made of two simpler conditions: "is a sectional", and "has a chaise". For simplicity, let's use p to designate "is a sectional", and q to designate "has a chaise".

A truth table for this situation would look like this:

p	q	p or q
T	T	T
T	F	T
F	T	T
F	F	F

In the table, T is used for true, and F for false. In the first row, if p is true and q is also true, then the complex statement "p or *q*" is true. This would be a sectional that also has a chaise, which meets our desire. (Remember that *or* in logic is not exclusive; if the couch has both features, it meets the condition.)

In the previous example about the couch, the truth table was really just summarizing what we already know about how the *or* statement work. The truth tables for the basic *and*, *or*, and *not* statements are shown below.

Basic Truth Tables

Negation - Expresses "not" which means the opposite truth value.

p	$\sim p$
T	F
F	T

Conjunction - Expresses "and" which means both p and q must be true.

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Disjunction - Expresses "or" which means either p or q can be true or both.

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Truth tables really become useful when we analyze more complex statements.

✔ Example 2

Create a truth table for the statement $p \vee \sim q$

Solution

When we create the truth table, we need to list all the possible truth value combinations for p and q. Notice how the first column contains 2 Ts followed by 2 Fs, and the second column alternates $\mathrm{T}, \mathrm{F}, \mathrm{T}$, F. This pattern ensures that all 4 combinations are considered.

p	q
T	T
T	F
F	T
F	F

After creating columns with those initial values, we create a third column for the expression $\sim q$. Now we will temporarily ignore the column for p and write the truth values for $\sim q$

p	q	$\sim q$
T	T	F
T	F	T
F	T	F
F	F	T

Next we can find the truth values of $p \vee \sim q$, using the first and third columns.

p	q	$\sim q$	$p \vee \sim q$
T	T	F	T
T	F	T	T
F	T	F	F
F	F	T	T

The truth table shows that $p \vee \sim q$ is true in three cases and false in one case. If you're wondering what the point of this is, suppose it is the last day of the baseball season and two teams, who are not playing each other, are competing for the final playoff spot. Anaheim will make the playoffs if it wins its game <u>or</u> if Boston does not win its game. (Anaheim owns the tie-breaker; if both teams win, or if both teams lose, then Anaheim gets the playoff spot.) If $p =$ Anaheim wins its game and $q =$ Boston wins its game, then $p \vee \sim q$ represents the situation "Anaheim wins its game or Boston does not win its game". The truth table shows us the different scenarios related to Anaheim making the playoffs. In the first row, Anaheim wins its game and Boston wins its game, so it is true that Anaheim makes the playoffs. In the second row, Anaheim wins and Boston does not win, so it is true that Anaheim makes the playoffs. In the third row, Anaheim does not win its game and Boston wins its game, so it is <u>false</u> that Anaheim makes the playoffs. In the fourth row, Anaheim does not win and Boston does not win, so it is true that Anaheim makes the playoffs.

? Try it Now 1

Create a truth table for this statement: $\sim p \wedge q$

Answer

p	q	$\sim p$	$\sim p \wedge q$
T	T	F	F
T	F	F	F
F	T	T	T
F	F	T	F

Example 3

Create a truth table for the statement $p \wedge \sim (q \vee r)$

Solution

It helps to work from the inside out when creating a truth table, and to create columns in the table for intermediate operations. We start by listing all the possible truth value combinations for $p, q,$ and r. Notice how the first column contains 4 Ts followed by $4\mathrm{F}$s, the second column contains $2\mathrm{T}$s, $2\mathrm{F}$s, then repeats, and the last column alternates $\mathrm{T}, \mathrm{F}, \mathrm{T}, \mathrm{F} \ldots$ This pattern ensures that all 8 combinations are considered. After creating columns with those initial values, we create a fourth column for the innermost expression, $q \vee r$. Now we will temporarily ignore the column for p and focus on q and r, writing the truth values for $q \vee r$

p	q	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

p	q	r	$q \vee r$
T	T	T	T
T	T	F	T
T	F	T	T
T	F	F	F
F	T	T	T
F	T	F	T
F	F	T	T
F	F	F	F

Next we can find the negation of $q \vee r$, working off the $q \vee r$ Column we just created. (Ignore the first three columns and simply negate the values in the $q \vee r$ column.)

p	q	r	$q \vee r$	$\sim (q \vee r)$
T	T	T	T	F
T	T	F	T	F
T	F	T	T	F
T	F	F	F	T
F	T	T	T	F
F	T	F	T	F
F	F	T	T	F
F	F	F	F	T

Finally, we find the values of p and $\sim (q \vee r)$. (Ignore the second, third, and fourth columns.)

p	q	r	$q \vee r$	$\sim(q \vee r)$	$p \wedge \sim(q \vee r)$
T	T	T	T	F	F
T	T	F	T	F	F
T	F	T	T	F	F
T	F	F	F	T	T
F	T	T	T	F	F
F	T	F	T	F	F
F	F	T	T	F	F
F	F	F	F	T	F

It turns out that this complex expression is true in only one case: when p is true, q is false, and r is false. To illustrate this situation, suppose that Anaheim will make the playoffs if: (1) Anaheim wins, and (2) neither Boston nor Cleveland wins. TFF is the only scenario in which Anaheim will make the playoffs.

Try it Now 2

Create a truth table for this statement: $(\sim p \wedge q) \vee \sim q$

Answer

p	q	$\sim p$	$\sim p \wedge q$	$\sim q$	$(\sim p \wedge q) \vee \sim q$
T	T	F	F	F	F
T	F	F	F	T	T
F	T	T	T	F	T
F	F	T	F	T	T

You have probably noticed that different truth tables have different numbers of rows, depending on how many variables (or simple statements) that we have. When there is only one simple statement (such as our truth table for negation), we have two rows, one for true and another for false. When we have two simple statements, p and q, there are four rows in the truth table. For Example 3, we had three simple statements, p, q and r, with eight rows in the truth table. So for each new simple statement, the number of rows doubles. This pattern continues, so we can generalize it as follows.

Definition: The Number of Lines in a Truth Table

A statement with k variables - or simple statements - will have a truth table with 2^k rows.

Logically Equivalent

Two statements are **logically equivalent** if they have the same simple statements and when their truth tables are computed, the final columns in the tables are identical.

DeMorgan's Laws

The negation of a conjunction is logically equivalent to the disjunction of the negation of the statements making up the conjunction. To negate an “and” statement, negate each part and change the “and” to “or”.

$$\sim(p \wedge q) \text{ is logically equivalent to } \sim p \vee \sim q$$

The negation of a disjunction is logically equivalent to the conjunction of the negation of the statements making up the disjunction. To negate an “or” statement, negate each part and change the “or” to “and”.

$$\sim(p \vee q) \text{ is logically equivalent to } \sim p \wedge \sim q$$

✔ Example 4

For Valentine's Day, you did not get your sweetie flowers or candy: Which of the following statements is logically equivalent?

a. You did not get them flowers or did not get them candy.
b. You did not get them flowers and did not get them candy.
c. You got them flowers or got them candy.

Solution

a. This statement does not go far enough; it leaves open the possibility that you got them one of the two things.
b. This statement is equivalent to the original; $\sim(f \vee c)$ is equivalent to $\sim f \wedge \sim c$
c. This statement says that you got them something, but we know that you did not.

? Try it Now 3

To serve as the President of the US, a person must have been born in the US, must be at least 35 years old, and must have lived in the US for at least 14 years. What minimum set of conditions would disqualify someone from serving as President?

Answer

Failing to meet just one of the three conditions is all it takes to be disqualified. A person is disqualified if they were not born in the US, or are not at least 35 years old, or have not lived in the US for at least 14 years. The key word here is "or" instead of "and".

3.3: Truth Tables- Conditional, Biconditional

Conditional

A **conditional** is a logical compound statement in which a statement p, called the hypothesis, implies a statement q, called the conclusion.

A conditional is written as $p \rightarrow q$ and is translated as "if p, then q".

Example 1

The English statement “If it is raining, then there are clouds is the sky” is a conditional statement. It makes sense because if the hypothesis “it is raining” is true, then the conclusion “there are clouds in the sky” must also be true.

Notice that the statement tells us nothing of what to expect if it is not raining; there might be clouds in the sky, or there might not. If the hypothesis is false, then the conclusion becomes irrelevant.

Example 2

Suppose you order a team jersey online on Tuesday and want to receive it by Friday so you can wear it to Saturday’s game. The website says that if you pay for expedited shipping, you will receive the jersey by Friday. In what situation is the website telling a lie?

There are four possible outcomes:

1. You pay for expedited shipping and receive the jersey by Friday
2. You pay for expedited shipping and don’t receive the jersey by Friday
3. You don’t pay for expedited shipping and receive the jersey by Friday
4. You don’t pay for expedited shipping and don’t receive the jersey by Friday

Only one of these outcomes proves that the website was lying: the second outcome in which you pay for expedited shipping but don’t receive the jersey by Friday. The first outcome is exactly what was promised, so there’s no problem with that. The third outcome is not a lie because the website never said what would happen if you didn’t pay for expedited shipping; maybe the jersey would arrive by Friday whether you paid for expedited shipping or not. The fourth outcome is not a lie because, again, the website didn’t make any promises about when the jersey would arrive if you didn’t pay for expedited shipping.

It may seem strange that the third outcome in the previous example, in which the first part is false but the second part is true, is not a lie. Remember, though, that if the hypothesis is false, we cannot make any judgment about the conclusion. The website never said that paying for expedited shipping was the *only* way to receive the jersey by Friday.

Example 3

A friend tells you “If you upload that picture to Facebook, you’ll lose your job.” Under what conditions can you say that your friend was wrong?

There are four possible outcomes:

1. You upload the picture and lose your job
2. You upload the picture and don’t lose your job
3. You don’t upload the picture and lose your job
4. You don’t upload the picture and don’t lose your job

There is only one possible case in which you can say your friend was wrong: the second outcome in which you upload the picture but still keep your job. In the last two cases, your friend didn’t say anything about what would happen if you didn’t upload the picture, so you can’t say that their statement was wrong. Even if you didn’t upload the picture and lost your job anyway, your friend never said that you were guaranteed to keep your job if you didn’t upload the picture; you might lose your job for missing a shift or punching your boss instead.

Truth Table for the Conditional

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

Again, if the hypothesis p is false, we cannot prove that the statement is a lie, so the result of the third and fourth rows is true.

Example 4

Construct a truth table for the statement $(m \wedge \sim p) \rightarrow r$

Solution

We start by constructing a truth table with 8 rows to cover all possible scenarios. Next, we can focus on the hypothesis, $m \wedge \sim p$.

m	p	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

m	p	r	$\sim p$
T	T	T	F
T	T	F	F
T	F	T	T
T	F	F	T
F	T	T	F
F	T	F	F
F	F	T	T
F	F	F	T

m	p	r	$\sim p$	$m \wedge \sim p$
T	T	T	F	F
T	T	F	F	F
T	F	T	T	T
T	F	F	T	T
F	T	T	F	F
F	T	F	F	F
F	F	T	T	F
F	F	F	T	F

Now we can create a column for the conditional. Because it can be confusing to keep track of all the Ts and Fs, why don't we copy the column for r to the right of the column for $m \wedge \sim p$? This makes it a lot easier to read the conditional from left to right.

m	p	r	$\sim p$	$m \wedge \sim p$	r	$(m \wedge \sim p) \rightarrow r$
T	T	T	F	F	T	T
T	T	F	F	F	F	T
T	F	T	T	T	T	T
T	F	F	T	T	F	F
F	T	T	F	F	T	T
F	T	F	F	F	F	T
F	F	T	T	F	T	T
F	F	F	T	F	F	T

When m is true, p is false, and r is false- -the fourth row of the table-then the hypothesis $m \wedge \sim p$ will be true but the conclusion false, resulting in an invalid conditional; every other case gives a valid conditional.

If you want a real-life situation that could be modeled by $(m \wedge \sim p) \rightarrow r$, consider this: let $m =$ we order meatballs, $p =$ we order pasta, and $r =$ Rob is happy. The statement $(m \wedge \sim p) \rightarrow r$ is "if we order meatballs and don't order pasta, then Rob is happy". If m is true (we order meatballs), p is false (we don't order pasta), and r is false (Rob is not happy), then the statement is false, because we satisfied the hypothesis but Rob did not satisfy the conclusion.

For any conditional, there are three related statements, the converse, the inverse, and the contrapositive.

Derived Forms of a Conditional

The original conditional is "if p, then q'' $p \rightarrow q$

The converse is "if q, then p'' $q \rightarrow p$

The inverse is "if not p, then not q'' $\sim p \rightarrow \sim q$

The contrapositive is "if not q, then not p'' $\sim q \rightarrow \sim p$

Example 5

Consider again the conditional “If it is raining, then there are clouds in the sky.” It seems reasonable to assume that this is true.

The converse would be “If there are clouds in the sky, then it is raining.” This is not always true.

The inverse would be “If it is not raining, then there are not clouds in the sky.” Likewise, this is not always true.

The contrapositive would be “If there are not clouds in the sky, then it is not raining.” This statement is true, and is equivalent to the original conditional.

Looking at truth tables, we can see that the original conditional and the contrapositive are logically equivalent, and that the converse and inverse are logically equivalent.

		Conditional	Converse	Inverse	Contrapositive
p	q	$p \rightarrow q$	$q \rightarrow p$	$\sim p \rightarrow \sim q$	$\sim q \rightarrow \sim p$
T	T	T	T	T	T
T	F	F	T	T	F
F	T	T	F	F	T
F	F	T	T	T	T

Equivalent

Equivalence

A conditional statement and its contrapositive are logically equivalent.

The converse and inverse of a conditional statement are logically equivalent.

In other words, the original statement and the contrapositive must agree with each other; they must both be true, or they must both be false. Similarly, the converse and the inverse must agree with each other; they must both be true, or they must both be false.

We typically represent the conditional using the words, "if ..., then ...," but there are other ways this logical connective can be represented in English. Consider the conditional from Example 5: "If it is raining, then there are clouds in the sky." We could equivalently write, "It is raining *only if* there are clouds in the sky." You can probably imagine how these two statements are saying the same thing - whenever it's raining outside, it is a safe conclusion there are clouds in the sky as well. Some other wordings that communicate the same information use either "sufficient" or "necessary." For example, "Raining is a sufficient condition for it to be cloudy," and "Being cloudy is a necessary condition for it to be raining." Here is a table summarizing the different wordings.

Different Wordings of the Conditional

The following statements are equivalent:

- If p, then q.
- q only if p.
- p is sufficient for q.
- q is necessary for p.

In everyday life, we often have a stronger meaning in mind when we use a conditional statement. Consider "If you submit your hours today, then you will be paid next Friday." What the payroll rep really means is "If you submit your hours today, then you will be paid next Friday, and if you don't submit your hours today, then you won't be paid next Friday." The conditional statement if *t*, then *p* also includes the inverse of the statement: if not *t*, then not *p*. A more compact way to express this statement is "You will be paid next Friday *if and only if* you submit your timesheet today." A statement of this form is called a **biconditional**.

Biconditional

A biconditional is a logical conditional statement in which the hypothesis and conclusion are interchangeable.

A biconditional is written as $p \leftrightarrow q$ and is translated as " p if and only if q''.

Because a biconditional statement $p \leftrightarrow q$ is equivalent to $(p \rightarrow q) \wedge (q \rightarrow p)$, we may think of it as a conditional statement combined with its converse: if p, then q and if q, then p. The double-headed arrow shows that the conditional statement goes from left to right and from right to left. A biconditional is considered true as long as the hypothesis and the conclusion have the same truth value; that is, they are either both true or both false.

Truth Table for the Biconditional

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

Notice that the fourth row, where both components are false, is true; if you don't submit your timesheet and you don't get paid, the person from payroll told you the truth.

Example 6

Suppose this statement is true: "The garbage truck comes down my street if and only if it is Thursday morning." Which of the following statements could be true?

a. It is noon on Thursday and the garbage truck did not come down my street this morning.
b. It is Monday and the garbage truck is coming down my street.
c. It is Wednesday at 11:59PM and the garbage truck did not come down my street today.

Solution

a. This cannot be true. This is like the second row of the truth table; it is true that I just experienced Thursday morning, but it is false that the garbage truck came.
b. This cannot be true. This is like the third row of the truth table; it is false that it is Thursday, but it is true that the garbage truck came.
c. This could be true. This is like the fourth row of the truth table; it is false that it is Thursday, but it is also false that the garbage truck came, so everything worked out like it should.

? Try it Now 1

Suppose this statement is true: "I wear my running shoes if and only if I am exercising." Determine whether each of the following statements must be true or false.

a. I am exercising and I am not wearing my running shoes.
b. I am wearing my running shoes and I am not exercising.
c. I am not exercising and I am not wearing my running shoes.

Answer

Choices a & b are false; c is true.

✔ Example 7

Create a truth table for the statement $(A \vee B) \leftrightarrow \sim C$

Solution

Whenever we have three component statements, we start by listing all the possible truth value combinations for A, B, and C. After creating those three columns, we can create a fourth column for the hypothesis, $A \vee B$. Now we will temporarily ignore the column for C and focus on A and B, writing the truth values for $A \vee B$.

A	B	C
T	T	T
T	T	F
T	F	T
T	F	F
F	T	T
F	T	F
F	F	T
F	F	F

A	B	C	$A \vee B$
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	F
F	F	F	F

Next we can create a column for the negation of C. (Ignore the $A \vee B$ column and simply negate the values in the C column.)

A	B	C	$A \vee B$	$\sim C$
T	T	T	T	F
T	T	F	T	T
T	F	T	T	F
T	F	F	T	T
F	T	T	T	F
F	T	F	T	T
F	F	T	F	F
F	F	F	F	T

Finally, we find the truth values of $(A \vee B) \leftrightarrow \sim C$. Remember, a biconditional is true when the truth value of the two parts match, but it is false when the truth values do not match.

A	B	C	$A \vee B$	$\sim C$	$(A \vee B) \leftrightarrow \sim C$
T	T	T	T	F	F
T	T	F	T	T	T
T	F	T	T	F	F
T	F	F	T	T	T
F	T	T	T	F	F
F	T	F	T	T	T
F	F	T	F	F	T
F	F	F	F	T	F

To illustrate this situation, suppose your boss needs you to do either project A or project B (or both, if you have the time). If you do one of the projects, you will not get a crummy review (C is for crummy). So $(A \vee B) \leftrightarrow \sim C$ means "You will not get a crummy review if and only if you do project A or project B." Looking at a few of the rows of the truth table, we can see how this works out. In the first row, A, B, and C are all true: you did both projects and got a crummy review, which is not what your boss told you would happen! That is why the final result of the first row is false. In the fourth row, A is true, B is false, and C is false: you did project A and did not get a crummy review. This is what your boss said would happen, so the final result of this row is true. And in the eighth row, A, B, and C are all false: you didn't do either project and did not get a crummy review. This is not what your boss said would happen, so the final result of this row is false. (Even though you may be happy that your boss didn't follow through on the threat, the truth table shows that your boss lied about what would happen.)

3.4: Verifying Arguments

Arguments

An **argument** is a series of statements called **premises** followed by a single statement called the **conclusion**.

An argument is **valid** if whenever all the premises are true, then the conclusion must also be true. Otherwise the argument is **invalid**.

Arguments can also be analyzed using truth tables, although this can be a lot of work.

Analyzing Arguments Using Truth Tables

To analyze an argument with a truth table:

1. Make a truth table with a seperate column for each premise and the conclusion.
2. Examine *only* the lines in the table which all of the premises are true.
3. If the conclusion is also true for the lines you examined in step 2, the argument is valid.
4. If the conclusion is false for even one of the lines you examined in step 2, the argument is invalid.

Example 1

Consider the argument

Premise:	If you bought bread, then you went to the store.
Premise:	You bought bread.
Conclusion:	You went to the store.

Solution

While this example is fairly obviously a valid argument, we can analyze it using a truth table by representing each of the premises symbolically.

We'll let b represent "you bought bread" and s represent "you went to the store". Then the argument becomes:

Premise:	$b \rightarrow s$
Premise:	b
Conclusion:	s

To test the validity, we create a truth table and a separate column for both premises and the conclusion:

b	s	$b \rightarrow s$	b	s
T	T	T	T	T
T	F	F	T	
F	T	T	F	
F	F	T	F	

We only examine the first line since that is the only time both premises are true. We see that the conclusion is also true. This argument is valid.

Example 2

Determine whether the argument is valid:

Premise:	If I have a shovel, I can dig a hole.
Premise:	I dug a hole.
Conclusion:	Therefore, I had a shovel.

Solution

Let $S =$ have a shovel, $D =$ dig a hole. The argument becomes:

Premise:	$S \rightarrow D$
Premise:	D
Conclusion:	S

We can construct a truth table with a seperate column for each premise and the conclusion:

S	D	$S \rightarrow D$	D	S
T	T	T	T	T
T	F	F	F	
F	T	T	T	F
F	F	T	F	

We need to examine both Line 1 and Line 3 since both premises are true in both cases. Since the conclusion is false in Line 3, the argument is invalid.

Rather than making a truth table for every argument, we may be able to recognize certain common forms of arguments that are valid (or invalid). If we can determine that an argument fits one of the common forms, we can immediately state whether it is valid or invalid.

The Law of Detachment

The law of detachment is:

Premise:	$p \rightarrow q$
Premise:	p
Conclusion:	q

Example 3

Recall this argument from an earlier example:

Premise:	If you bought bread, then you went to the store.
Premise:	You bought bread.
Conclusion:	You went to the store.

In symbolic form:

Premise:	$b \rightarrow s$
Premise:	b
Conclusion:	s

This argument has the structure described by the law of detachment. (The second premise and the conclusion are simply the two parts of the first premise detached from each other.) Instead of making a truth table, we can say that this argument is valid by stating that it satisfies the law of detachment.

The Law of Contraposition

The law of contraposition is:

Premise:	$p \rightarrow q$
Premise:	$\sim q$
Conclusion:	$\sim p$

Notice that the second premise and the conclusion look like the contrapositive of the first premise, $\sim q \rightarrow \sim p$, but they have been detached. You can think of the law of contraposition as a combination of the law of detachment and the fact that the contrapositive is logically equivalent to the original statement.

Example 4

Premise:	If I drop my phone into the swimming pool, my phone will be ruined.
Premise:	My phone isn't ruined.
Conclusion:	I didn't drop my phone into the swimming pool.

If we let $d = \text{I}$ drop the phone in the pool and $r =$ the phone is ruined, then we can represent the argument this way:

Premise:	$d \to r$
Premise:	$\sim r$
Conclusion:	$\sim d$

The form of this argument matches what we need to invoke the law of contraposition, so it is a valid argument.

Try it Now 1

Is this argument valid?

Premise:	If you brushed your teeth before bed, then your toothbrush will be wet.
Premise:	Your toothbrush is dry.
Conclusion:	You didn't brush your teeth before bed.

Answer

Let $b =$ brushed teeth and $w =$ toothbrush is wet.

Premise:	$b \to w$
Premise:	$\sim w$
Conclusion:	$\sim b$

This argument is valid by the Law of Contraposition.

The Law of Syllogism

The Law of Syllogism is:

Premise:	$p \to q$
Premise:	$q \to r$
Conclusion:	$p \to r$

The earlier example about buying a shirt at the mall is an example illustrating the transitive property. It describes a chain reaction: if the first thing happens, then the second thing happens, and if the second thing happens, then the third thing happens. Therefore, if we want to ignore the second thing, we can say that if the first thing happens, then we know the third thing will happen. We don't have to mention the part about buying jeans; we can simply say that the first event leads to the final event. We could even have more than two premises; as long as they form a chain reaction, the transitive property will give us a valid argument.

Example 5

Premise:	If a soccer player commits a reckless foul, she will receive a yellow card.
Premise:	If Hayley receives a yellow card, she will be suspended for the next match.
Conclusion:	If Hayley commits a reckless foul, she will be suspended for the next match.

If we let $r =$ committing a reckless foul, $y =$ receiving a yellow card, and $s =$ being suspended, then our argument looks like this:

Premise:	$r \to y$
Premise:	$y \to s$
Conclusion:	$r \to s$

This argument has the exact structure required to use the transitive property, so it is a valid argument.

Try it Now 2

Is this argument valid?

Premise:	If the old lady swallows a fly, she will swallow a spider.
Premise:	If the old lady swallows a spider, she will swallow a bird.
Premise:	Premise: If the old lady swallows a bird, she will swallow a cat.
Premise:	If the old lady swallows a cat, she will swallow a dog.
Premise:	If the old lady swallows a dog, she will swallow a goat.
Premise:	If the old lady swallows a goat, she will swallow a cow.
Premise:	If the old lady swallows a cow, she will swallow a horse.
Premise:	If the old lady swallows a horse, she will die, of course.
Conclusion:	If the old lady swallows a fly, she will die, of course.

Answer

This argument is valid by the Law of Syllogism, which can involve more than two premises, as long as they continue the chain reaction. The premises $f \rightarrow s, s \rightarrow b, b \rightarrow c, c \rightarrow d \quad d \rightarrow g, g \rightarrow w, w \rightarrow h, h \rightarrow x$ can be reduced to $f \rightarrow x$. (Because we had already used c and d we decided to use w for cow and x for death. If the old lady swallows the fly, she will eventually eat a horse and die.

Disjunctive Syllogism

Disjunctive syllogism is:

Premise:	$p \vee q$
Premise:	$\sim p$
Conclusion:	q

The order of the two parts of the disjunction isn't important. In other words, we could have the premises $p \vee q$ and $\sim q$, and the conclusion p

Example 6

Premise:	I can either drive or take the train.
Premise:	I refuse to drive.
Conclusion:	I will take the train.

If we let $d = I$ drive and $t = I$ take the train, then the symbolic representation of the argument is:

Premise:	$d \vee t$
Premise:	$\sim d$
Conclusion:	t

This argument is valid because it has the form of a disjunctive syllogism. I have two choices, and one of them is not going to happen, so the other one must happen.

Try it Now 3

Is this argument valid?

Premise:	Alison was required to write a 10-page paper or give a 5-minute speech.
Premise:	Alison did not give a 5-minute speech.
Conclusion:	Alison wrote a 10-page paper.

Answer

Let $p =$ wrote a paper and $s =$ gave a speech.

Premise:	$p \vee s$
Premise:	$-s$
Conclusion:	p

This argument is valid by Disjunctive Syllogism. Alison had to do one or the other; she didn't choose the speech, so she must have chosen the paper.

Keep in mind that, when you are determining the validity of an argument, you must assume that the premises are true. If you don't agree with one of the premises, you need to keep your personal opinion out of it. Your job is to pretend that the premises are true and then determine whether they force you to accept the conclusion. You may attack the premises in a court of law or a political discussion, of course, but here we are focusing on the structure of the arguments, not the truth of what they actually say.

We have just looked at four forms of valid arguments; there are two common forms that represent *invalid* arguments, which are also called *fallacies*.

The Fallacy of the Converse

The fallacy of the converse is:

Premise:	$p \rightarrow q$
Premise:	q
Conclusion:	p

Notice that the second premise and the conclusion look like the converse of the first premise, $q \rightarrow p$, but they have been detached. The fallacy of the converse incorrectly tries to assert that the converse of a statement is equivalent to that statement.

Example 7

Premise:	If I drink coffee after noon, then I have a hard time falling asleep that night.
Premise:	I had a hard time falling asleep last night.
Conclusion:	I drank coffee after noon yesterday.

If we let $c = \text{I}$ drink coffee after noon and $h = \text{I}$ have a hard time falling asleep, then our argument looks like this:

Premise:	$c \rightarrow h$
Premise:	h
Conclusion:	c

This argument uses converse reasoning, so it is an invalid argument. There could be plenty of other reasons why I couldn't fall asleep: I could be worried about money, my neighbors might have been setting off fireworks, …

Try it Now 4

Is this argument valid?

Premise:	If you pull that fire alarm, you will get in big trouble.
Premise:	You got in big trouble.
Conclusion:	You must have pulled the fire alarm.

Answer

Let $f =$ pulled fire alarm and $t =$ got in big trouble.

Premise:	$f \rightarrow t$
Premise:	t
Conclusion:	f

The Fallacy of the Inverse

The fallacy of the inverse is:

Premise:	$p \rightarrow q$
Premise:	$\sim p$
Conclusion:	$\sim q$

Again, notice that the second premise and the conclusion look like the inverse of the first premise, $\sim p \rightarrow \sim q$, but they have been detached. The fallacy of the inverse incorrectly tries to assert that the inverse of a statement is equivalent to that statement.

Example 8

Premise: If you listen to the Grateful Dead, then you are a hippie.
Premise: Sky doesn't listen to the Grateful Dead.
Conclusion: Sky is not a hippie.

If we let $g =$ listen to the Grateful Dead and $h =$ is a hippie, then this is the argument:

Premise: $g \rightarrow h$
Premise: $\sim g$
Conclusion: $\sim h$

This argument is invalid because it uses inverse reasoning. The first premise does not imply that all hippies listen to the Grateful Dead; there could be some hippies who listen to Phish instead.

Try it Now 5

Is this argument valid?

Premise: If a hockey player trips an opponent, he will be assessed a 2-minute penalty.
Premise: Alexei did not trip an opponent.
Conclusion: Alexei will not be assessed a 2-minute penalty.

Answer

Let $t =$ tripped and $p =$ got a penalty.

Premise: $t \rightarrow p$
Premise: $\sim t$
Conclusion: $\sim p$

This argument is invalid because it has the form of the Fallacy of the Inverse. Alexei may have gotten a penalty for an infraction other than tripping.

Of course, arguments are not limited to these six basic forms; some arguments have more premises, or premises that need to be rearranged before you can see what is really happening. There are plenty of other forms of arguments that are valid or invalid. If an argument doesn't seem to fit the pattern of any of these common forms, though, you may want to use a truth table instead.

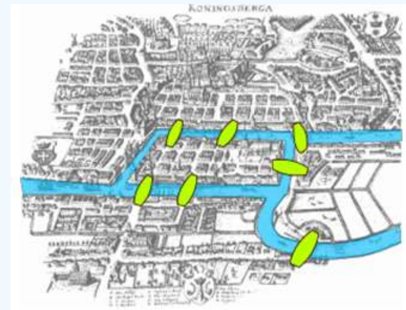
CHAPTER OVERVIEW

4: Graph Theory

4.1: Graphs

✔ Example 1

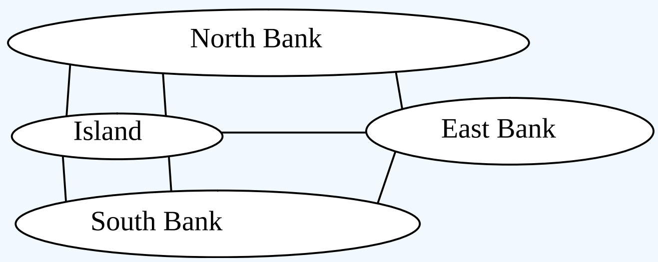
Back in the 18th century in the Prussian city of Königsberg, a river ran through the city and seven bridges crossed the forks of the river. The river and the bridges are highlighted in the picture to the right[1].



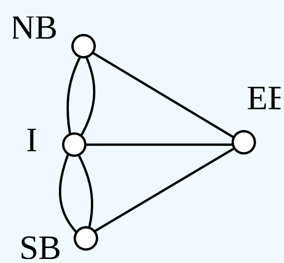
As a weekend amusement, townsfolk would see if they could find a route that would take them across every bridge once and return them to where they started.

Solution

Leonard Euler (pronounced OY-lur), one of the most prolific mathematicians ever, looked at this problem in 1735, laying the foundation for graph theory as a field in mathematics. To analyze this problem, Euler introduced edges representing the bridges:



Since the size of each land mass it is not relevant to the question of bridge crossings, each can be shrunk down to a vertex representing the location:



Notice that in this graph there are *two* edges connecting the north bank and island, corresponding to the two bridges in the original drawing. Depending upon the interpretation of edges and vertices appropriate to a scenario, it is entirely possible and reasonable to have more than one edge connecting two vertices.

While we haven't answered the actual question yet of whether or not there is a route which crosses every bridge once and returns to the starting location, the graph provides the foundation for exploring this question.

Definitions

While we loosely defined some terminology earlier, we now will try to be more specific.

Vertex

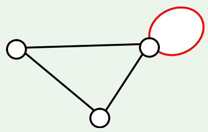
A **vertex** is a dot in the graph that could represent an intersection of streets, a land mass, or a general location, like "work" or "school". Vertices are often connected by edges. Note that vertices only occur when a dot is explicitly placed, not whenever two edges cross. Imagine a freeway overpass – the freeway and side street cross, but it is not possible to change from the side street to the freeway at that point, so there is no intersection and no vertex would be placed.

Edge

Edges connect pairs of vertices. An edge can represent a physical connection between locations, like a street, or simply that a route connecting the two locations exists, like an airline flight.

Loop

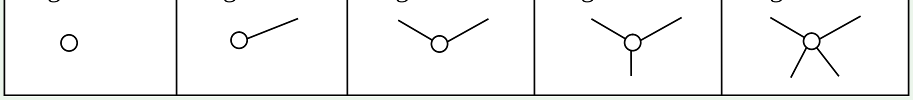
A **loop** is a special type of edge that connects a vertex to itself. Loops are not used much in street network graphs.



Degree of a Vertex

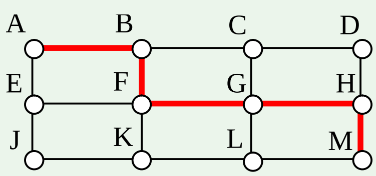
The **degree** of a vertex is the number of edges meeting at that vertex. It is possible for a vertex to have a degree of zero or larger.

Degree 0	Degree 1	Degree2	Degree 3	Degree 4



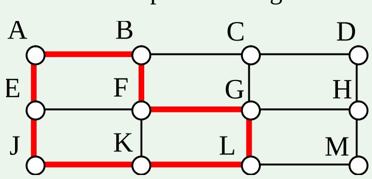
Path

A **path** is a sequence of vertices using the edges. Usually we are interested in a path between two vertices. For example, a path from vertex A to vertex M is shown below. It is one of many possible paths in this graph.



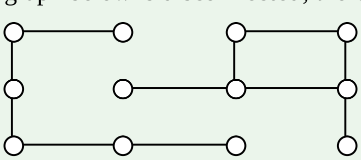
Circuit

A **circuit** is a path that begins and ends at the same vertex. A circuit starting and ending at vertex A is shown below.



Connected

A graph is **connected** if there is a path from any vertex to any other vertex. Every graph drawn so far has been connected. The graph below is **disconnected**; there is no way to get from the vertices on the left to the vertices on the right.



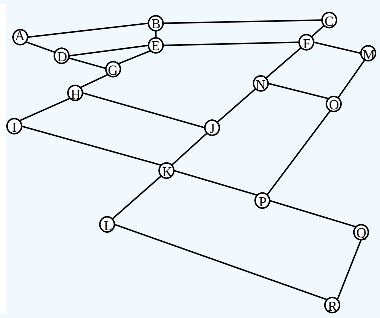
✓ Example 2

Here is a portion of a housing development from Missoula, Montana[2]. As part of her job, the development's lawn inspector has to walk down every street in the development making sure homeowners' landscaping conforms to the community requirements.

Solution

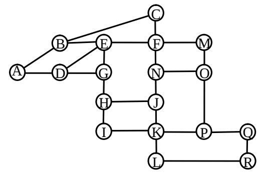
Naturally, she wants to minimize the amount of walking she has to do. Is it possible for her to walk down every street in this development without having to do any backtracking? While you might be able to answer that question just by looking at the picture for a while, it would be ideal to be able to answer the question for any picture regardless of its complexity.

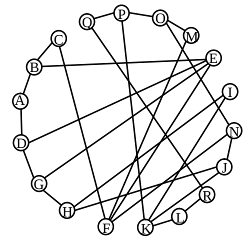
To do that, we first need to simplify the picture into a form that is easier to work with. We can do that by drawing a simple line for each street. Where streets intersect, we will place a dot.



This type of simplified picture is called a **graph**.

While we drew our original graph to correspond with the picture we had, there is nothing particularly important about the layout when we analyze a graph. Both of the graphs below are equivalent to the one drawn above since they show the same edge connections between the same vertices as the original graph.





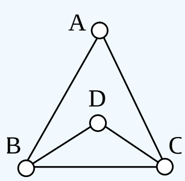
In the first section, we created a graph of the Königsberg bridges and asked whether it was possible to walk across every bridge once. Because Euler first studied this question, these types of paths are named after him.

Euler Path

An **Euler path** is a path that uses every edge in a graph with no repeats. Being a path, it does not have to return to the starting vertex.

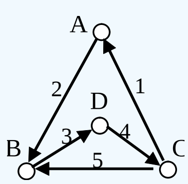
Example 4

In the graph shown below, there are several Euler paths.



Solution

One such path is CABDCB. The path is shown in arrows to the right, with the order of edges numbered.

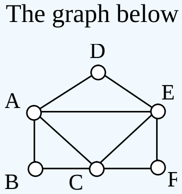


Euler Circuit

An **Euler circuit** is a circuit that uses every edge in a graph with no repeats. Being a circuit, it must start and end at the same vertex.

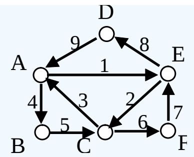
Example 5

The graph below has several possible Euler circuits.



Solution

Here's a couple, starting and ending at vertex A: ADEACEFCBA and AECABCFEDA. The second is shown in arrows.



Look back at the example used for Euler paths – does that graph have an Euler circuit? A few tries will tell you no; that graph does not have an Euler circuit. When we were working with shortest paths, we were interested in the optimal path. With Euler paths and circuits, we're primarily interested in whether an Euler path or circuit *exists*.

Why do we care if an Euler circuit exists? Think back to our housing development lawn inspector from the beginning of the chapter. The lawn inspector is interested in walking as little as possible. The ideal situation would be a circuit that covers every street with no repeats. That's an Euler circuit! Luckily, Euler solved the question of whether or not an Euler path or circuit will exist.

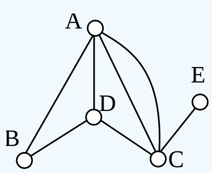
Euler's Theorem

A graph will contain an Euler path if it contains at most two vertices of odd degree.

A graph will contain an Euler circuit if all vertices have even degree

Example 6

In the graph below, vertices A and C have degree 4, since there are 4 edges leading into each vertex. B is degree 2, D is degree 3, and E is degree 1. This graph contains two vertices with odd degree (D and E) and three vertices with even degree (A, B, and C), so Euler's theorems tell us this graph has an Euler path, but not an Euler circuit.



Now we know how to determine if a graph has an Euler circuit, but if it does, how do we find one? While it usually is possible to find an Euler circuit just by pulling out your pencil and trying to find one, the more formal method is **Fleury's algorithm.**

Fleury's Algorithm

1. Start at any vertex if finding an Euler circuit. If finding an Euler path, start at one of the two vertices with odd degree.
2. Choose any edge leaving your current vertex, provided deleting that edge will not separate the graph into two disconnected sets of edges.
3. Add that edge to your circuit, and delete it from the graph.
4. Continue until you're done.

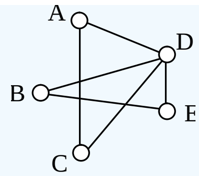
Example 7

Let's find an Euler Circuit on this graph using Fleury's algorithm, starting at vertex A.

Solution

Step 1:

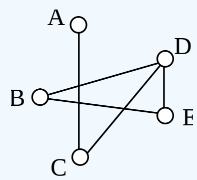
Original Graph. Choosing edge AD.



Circuit so far: AD

Step 2:

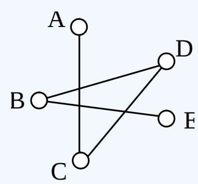
AD deleted. D is current. Can't choose DC since that would disconnect graph. Choosing DE



Circuit so far: ADE

Step 3:

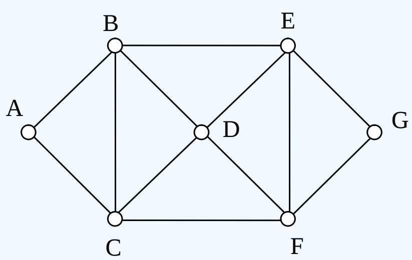
E is current. From here, there is only one option, so the rest of the circuit is determined.



Circuit: ADEBDCA

Try it Now 1

Does the graph below have an Euler Circuit? If so, find one.



Answer

Yes, all vertices have even degree so this graph has an Euler Circuit. There are several possibilities. One is: ABEGFCDFEDBCA

Not every graph has an Euler path or circuit, yet our lawn inspector still needs to do her inspections. Her goal is to minimize the amount of walking she has to do. In order to do that, she will have to duplicate some edges in the graph until an Euler circuit exists.

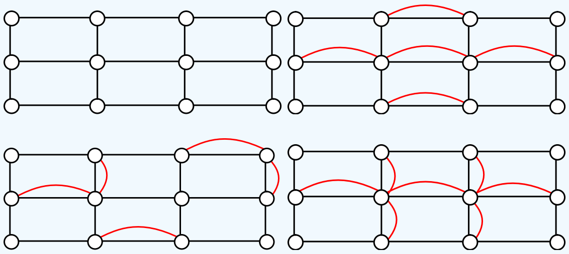
Eulerization

Eulerization is the process of adding edges to a graph to create an Euler circuit on a graph. To eulerize a graph, edges are duplicated to connect pairs of vertices with odd degree. Connecting two odd degree vertices increases the degree of each, giving them both even degree. When two odd degree vertices are not directly connected, we can duplicate all edges in a path connecting the two.

Note that we can only duplicate edges, not create edges where there wasn't one before. Duplicating edges would mean walking or driving down a road twice, while creating an edge where there wasn't one before is akin to installing a new road!

Example 8

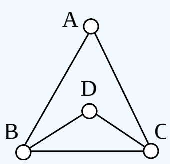
For the rectangular graph shown, three possible eulerizations are shown. Notice in each of these cases the vertices that started with odd degrees have even degrees after eulerization, allowing for an Euler circuit.



In the example above, you'll notice that the last eulerization required duplicating seven edges, while the first two only required duplicating five edges. If we were eulerizing the graph to find a walking path, we would want the eulerization with minimal duplications.

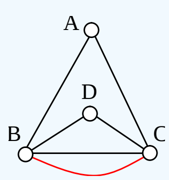
Try it Now 2

Eulerize the graph shown, then find an Euler circuit on the eulerized graph.



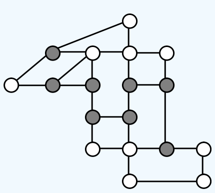
Answer

This graph can be eulerized by duplicating the edge BC, as shown. One possible Euler circuit on the eulerized graph is ACDBCBA



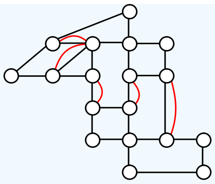
Example 9

Looking again at the graph for our lawn inspector, the vertices with odd degree are shown highlighted. With eight vertices, we will always have to duplicate at least four edges.



Solution

In this case, we need to duplicate five edges since two odd degree vertices are not directly connected. Without weights we can't be certain this is the eulerization that minimizes walking distance, but it looks pretty good.



The problem of finding the optimal eulerization is called the Chinese Postman Problem, a name given by an American in honor of the Chinese mathematician Mei-Ko Kwan who first studied the problem in 1962 while trying to find optimal delivery routes for postal carriers. This problem is important in determining efficient routes for garbage trucks, school buses, parking meter checkers, street sweepers, and more.

Unfortunately, algorithms to solve this problem are fairly complex.

---

[1] Bogdan Giuşcă. http://en.Wikipedia.org/wiki/File:Ko...rg_bridges.png

[2] Sam Beebe. http://www.flickr.com/photos/sbeebe/2850476641/, CC-BY

4.2: Hamiltonian Circuits and the Traveling Salesman Problem

In the last section, we considered optimizing a walking route for a postal carrier. How is this different than the requirements of a package delivery driver? While the postal carrier needed to walk down every street (edge) to deliver the mail, the package delivery driver instead needs to visit every one of a set of delivery locations. Instead of looking for a circuit that covers every edge once, the package deliverer is interested in a circuit that visits every vertex once.

Hamilton Circuits and Paths

A **Hamiltonian circuit** is a circuit that visits every vertex once with no repeats. Being a circuit, it must start and end at the same vertex. A **Hamiltonian path** also visits every vertex once with no repeats, but does not have to start and end at the same vertex.

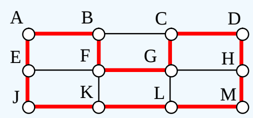
Hamiltonian circuits are named for William Rowan Hamilton who studied them in the 1800's.

Example 1

One Hamiltonian circuit is shown on the graph below. There are several other Hamiltonian circuits possible on this graph. Notice that the circuit only has to visit every vertex once; it does not need to use every edge.

Solution

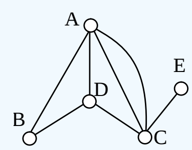
This circuit could be notated by the sequence of vertices visited, starting and ending at the same vertex: ABFGCDHMLKJEA. Notice that the same circuit could be written in reverse order, or starting and ending at a different vertex.



Unlike with Euler circuits, there is no nice theorem that allows us to instantly determine whether or not a Hamiltonian circuit exists for all graphs. [1]

Example 2

Does a Hamiltonian path or circuit exist on the graph below?



Solution

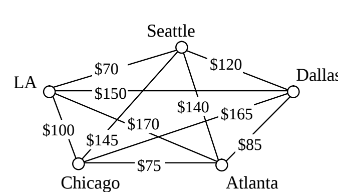
We can see that once we travel to vertex E there is no way to leave without returning to C, so there is no possibility of a Hamiltonian circuit. If we start at vertex E we can find several Hamiltonian paths, such as ECDAB and ECABD.

Weights

Depending on the problem being solved, sometimes weights are assigned to the edges. The weights could represent the distance between two locations, the travel time, or the travel cost.

With Hamiltonian circuits, our focus will not be on existence, but on the question of optimization; given a graph where the edges have weights, can we find the optimal Hamiltonian circuit; the one with lowest total weight.

This problem is called the **Traveling salesman problem** (TSP) because the question can be framed like this: Suppose a salesman needs to give sales pitches in four cities. He looks up the airfares between each city, and puts the costs in a graph. In what order should he travel to visit each city once then return home with the lowest cost?



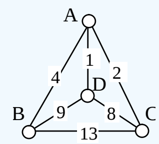
To answer this question of how to find the lowest cost Hamiltonian circuit, we will consider some possible approaches. The first option that might come to mind is to just try all different possible circuits.

Brute Force Algorithm (a.k.a. exhaustive search)

1. List all possible Hamiltonian circuits
2. Find the length of each circuit by adding the edge weights
3. Select the circuit with minimal total weight.

Example 3

Apply the Brute force algorithm to find the minimum cost Hamiltonian circuit on the graph below.



Solution

To apply the Brute force algorithm, we list all possible Hamiltonian circuits and calculate their weight:

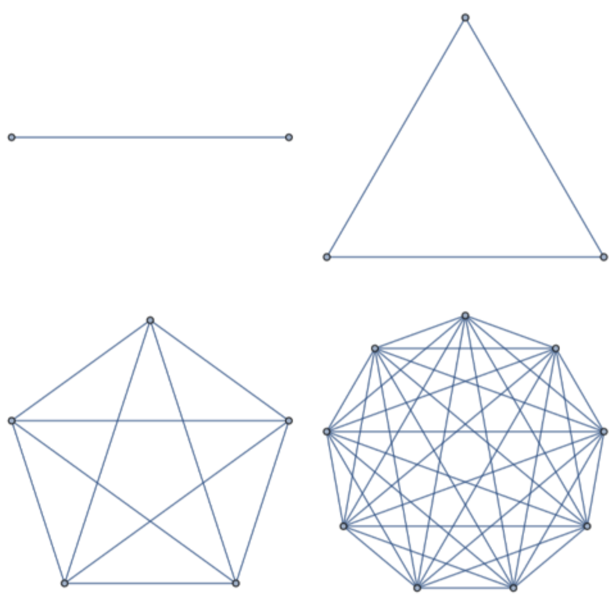
Circuit	Weight
ABCDA	$4+13+8+1=26$
ABDCA	$4+9+8+2=23$
ACBDA	$2+13+9+1=25$

Note: These are the unique circuits on this graph. All other possible circuits are the reverse of the listed ones or start at a different vertex, but result in the same weights.

From this we can see that the second circuit, ABDCA, is the optimal circuit.

The Brute force algorithm is optimal; it will always produce the Hamiltonian circuit with minimum weight. Is it efficient? To answer that question, we need to consider how many Hamiltonian circuits a graph could have. For simplicity, let's look at the worst-case possibility, where every vertex is connected to every other vertex. This is called a **complete graph.** In figure A, there are examples of complete graphs with different numbers of vertices.

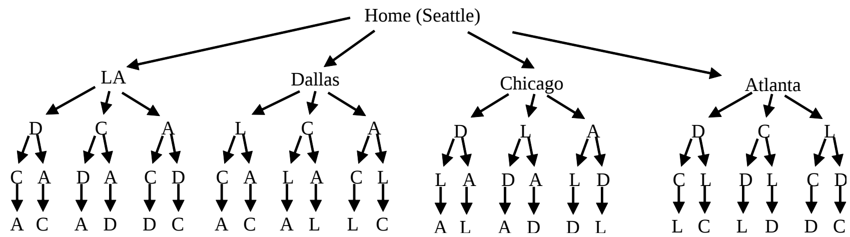
A



Out of convenience, mathematicians sometimes use specific notation for complete graphs based on the number of vertices. A complete graph with n vertices can be represented by K_n. For example, the graphs represented above are K_2, K_3, K_5 and K_9.

Suppose we had a complete graph with five vertices like the air travel graph above. From Seattle there are four cities we can visit first. From each of those, there are three choices. From each of those cities, there are two possible cities to visit next. There is then only one choice for the last city before returning home.

This can be shown visually:



Counting the number of routes, we can see there are $4 \cdot 3 \cdot 2 \cdot 1 = 24$ routes. For six cities there would be $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ routes.

Number of Possible Circuits

For n vertices in a complete graph, there will be $(n-1)! = (n-1)(n-2)(n-3)\cdots 3 \cdot 2 \cdot 1$ routes. Half of these are duplicates in reverse order, so there are $\frac{(n-1)!}{2}$ unique circuits.

The exclamation symbol, !, is read “factorial” and is shorthand for the product shown.

Example 4

How many circuits would a complete graph with 8 vertices have?

Solution

A complete graph with 8 vertices would have $(8-1)! = 7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$ possible Hamiltonian circuits. Half of the circuits are duplicates of other circuits but in reverse order, leaving 2520 unique routes.

While this is a lot, it doesn’t seem unreasonably huge. But consider what happens as the number of cities increase:

Cities	Unique Hamiltonian Circuits
9	$8!/2 = 20,160$
10	$9!/2 = 181,440$
11	$10!/2 = 1,814,400$
15	$14!/2 = 43,589,145,600$
20	$19!/2 = 60,822,550,204,416,000$

As you can see the number of circuits is growing extremely quickly. If a computer looked at one billion circuits a second, it would still take almost two years to examine all the possible circuits with only 20 cities! Certainly Brute Force is **not** an efficient algorithm.

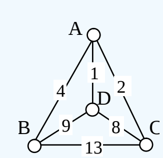
Unfortunately, no one has yet found an efficient *and* optimal algorithm to solve the TSP, and it is very unlikely anyone ever will. Since it is not practical to use brute force to solve the problem, we turn instead to **heuristic algorithms**; efficient algorithms that give approximate solutions. In other words, heuristic algorithms are fast, but may or may not produce the optimal circuit.

Nearest Neighbor Algorithm (NNA)

1. Select a starting point.
2. Move to the nearest unvisited vertex (the edge with smallest weight).
3. Repeat until the circuit is complete.

Example 5

Consider our earlier graph, shown to the right.



Starting at vertex A, the nearest neighbor is vertex D with a weight of 1.

From D, the nearest neighbor is C, with a weight of 8.

From C, our only option is to move to vertex B, the only unvisited vertex, with a cost of 13.

From B we return to A with a weight of 4.

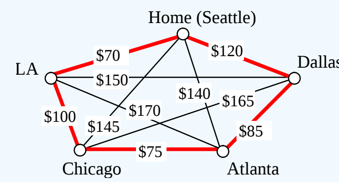
Solution

The resulting circuit is ADCBA with a total weight of $1 + 8 + 13 + 4 = 26$.

We ended up finding the worst circuit in the graph! What happened? Unfortunately, while it is very easy to implement, the NNA is a **greedy** algorithm, meaning it only looks at the immediate decision without considering the consequences in the future. In this case, following the edge AD forced us to use the very expensive edge BC later.

✔ Example 6

Consider again our salesman. Starting in Seattle, the nearest neighbor (cheapest flight) is to LA, at a cost of $70. From there:



Solution

- LA to Chicago: $100
- Chicago to Atlanta: $75
- Atlanta to Dallas: $85
- Dallas to Seattle: $120

Total cost: $450

In this case, nearest neighbor did find the optimal circuit.

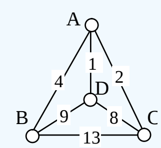
Going back to our first example, how could we improve the outcome? One option would be to redo the nearest neighbor algorithm with a different starting point to see if the result changed. Since nearest neighbor is so fast, doing it several times isn't a big deal.

✎ Repeated Nearest Neighbor Algorithm (RNNA)

1. Do the Nearest Neighbor Algorithm starting at each vertex
2. Choose the circuit produced with minimal total weight

✔ Example 7

We will revisit the graph from Example 17.



Starting at vertex A resulted in a circuit with weight 26.

Starting at vertex B, the nearest neighbor circuit is BADCB with a weight of 4+1+8+13 = 26. This is the same circuit we found starting at vertex A. No better.

Starting at vertex C, the nearest neighbor circuit is CADBC with a weight of 2+1+9+13 = 25. Better!

Starting at vertex D, the nearest neighbor circuit is DACBA. Notice that this is actually the same circuit we found starting at C, just written with a different starting vertex.

Solution

The RNNA was able to produce a slightly better circuit with a weight of 25, but still not the optimal circuit in this case. Notice that even though we found the circuit by starting at vertex C, we could still write the circuit starting at A: ADBCA or ACBDA.

? Try it Now 1

The table below shows the time, in milliseconds, it takes to send a packet of data between computers on a network. If data needed to be sent in sequence to each computer, then notification needed to come back to the original computer, we would be solving the TSP. The computers are labeled A-F for convenience.

	A	B	C	D	E	F
A	__	44	34	12	40	41
B	44	__	31	43	24	50
C	34	31	__	20	39	27
D	12	43	20	__	11	17
E	40	24	39	11	__	42
F	41	50	27	17	42	__

a. Find the circuit generated by the NNA starting at vertex B.

b. Find the circuit generated by the RNNA.

Answer

At each step, we look for the nearest location we haven’t already visited.

From B the nearest computer is E with time 24.

From E, the nearest computer is D with time 11.

From D the nearest is A with time 12.

From A the nearest is C with time 34.

From C, the only computer we haven’t visited is F with time 27

From F, we return back to B with time 50.

The NNA circuit from B is BEDACFB with time 158 milliseconds.

While certainly better than the basic NNA, unfortunately, the RNNA is still greedy and will produce very bad results for some graphs. As an alternative, our next approach will step back and look at the “big picture” – it will select first the edges that are shortest, and then fill in the gaps.

Cheapest Edge Algorithm (Best Edge/Greedy Algorithm)

1. Select the cheapest unused edge in the graph.
2. Repeat step 1, adding the cheapest unused edge to the circuit, unless:
    a. adding the edge would create a circuit that doesn’t contain all vertices, or
    b. adding the edge would give a vertex degree 3.
3. Repeat until a circuit containing all vertices is formed.

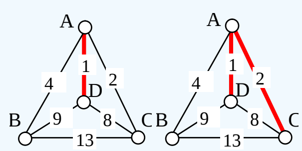
Example 8

Using the four vertex graph from earlier, we can use the Sorted Edges algorithm.

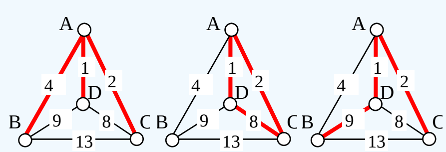
Solution

The cheapest edge is AD, with a cost of 1. We highlight that edge to mark it selected.

The next shortest edge is AC, with a weight of 2, so we highlight that edge.

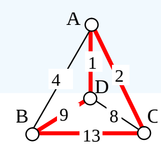


For the third edge, we’d like to add AB, but that would give vertex A degree 3, which is not allowed in a Hamiltonian circuit. The next shortest edge is CD, but that edge would create a circuit ACDA that does not include vertex B, so we reject that edge. The next shortest edge is BD, so we add that edge to the graph.



We then add the last edge to complete the circuit: ACBDA with weight 25.

Notice that the algorithm did not produce the optimal circuit in this case; the optimal circuit is ACDBA with weight 23.



While the Sorted Edge algorithm overcomes some of the shortcomings of NNA, it is still only a heuristic algorithm, and does not guarantee the optimal circuit.

Example 9

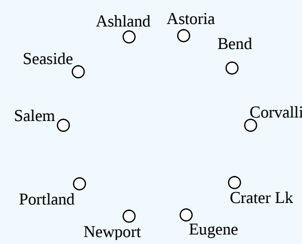
Your teacher’s band, *Derivative Work*, is doing a bar tour in Oregon. The driving distances are shown below. Plan an efficient route for your teacher to visit all the cities and return to the starting location. Use NNA starting at Portland, and then use Sorted Edges.

	Ashland	Astoria	Bend	Corvallis	Crater Lake	Eugene	Newport	Portland	Salem	Seaside
Ashland	_	374	200	223	108	178	252	285	240	356
Astoria	374	_	255	166	433	199	135	95	136	17
Bend	200	255	_	128	277	128	180	160	131	247
Corvallis	223	166	128	_	430	47	52	84	40	155
Crater Lake	108	433	277	430	_	453	478	344	389	423
Eugene	178	199	128	47	453	_	91	110	64	181
Newport	252	135	180	52	478	91	_	114	83	117
Portland	285	95	160	84	344	110	114	_	47	78
Salem	240	136	131	40	389	64	83	47	_	118
Seaside	356	17	247	155	423	181	117	78	118	_

Solution

Using NNA with a large number of cities, you might find it helpful to mark off the cities as they're visited to keep from accidently visiting them again. Looking in the row for Portland, the smallest distance is 47, to Salem. Following that idea, our circuit will be:

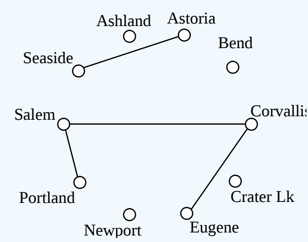
Portland to Salem 47
Salem to Corvallis 40
Corvallis to Eugene 47
Eugene to Newport 91
Newport to Seaside 117
Seaside to Astoria 17
Astoria to Bend 255
Bend to Ashland 200
Ashland to Crater Lake 108
Crater Lake to Portland 344
Total trip length: 1266 miles



Using Sorted Edges, you might find it helpful to draw an empty graph, perhaps by drawing vertices in a circular pattern. Adding edges to the graph as you select them will help you visualize any circuits or vertices with degree 3.

We start adding the shortest edges:

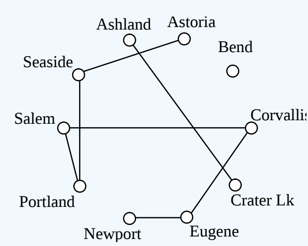
Seaside to Astoria 17 miles
Corvallis to Salem 40 miles
Portland to Salem 47 miles
Corvallis to Eugene 47 miles



The graph after adding these edges is shown to the right. The next shortest edge is from Corvallis to Newport at 52 miles, but adding that edge would give Corvallis degree 3.

Continuing on, we can skip over any edge pair that contains Salem or Corvallis, since they both already have degree 2.

Portland to Seaside 78 miles
Eugene to Newport 91 miles
Portland to Astoria (reject – closes circuit)
Ashland to Crater Lk 108 miles



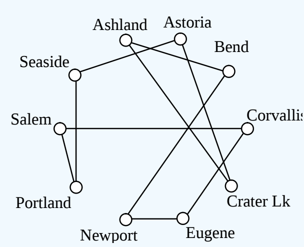
The graph after adding these edges is shown to the right. At this point, we can skip over any edge pair that contains Salem, Seaside, Eugene, Portland, or Corvallis since they already have degree 2.

Newport to Astoria (reject – closes circuit)
Newport to Bend 180 miles
Bend to Ashland 200 miles

At this point the only way to complete the circuit is to add:

Crater Lk to Astoria 433 miles

The final circuit, written to start at Portland, is:



Portland, Salem, Corvallis, Eugene, Newport, Bend, Ashland, Crater Lake, Astoria, Seaside, Portland.

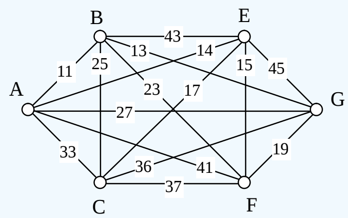
Total trip length: 1241 miles.

While better than the NNA route, neither algorithm produced the optimal route. The following route can make the tour in 1069 miles:

Portland, Astoria, Seaside, Newport, Corvallis, Eugene, Ashland, Crater Lake, Bend, Salem, Portland

? Try it Now 2

Find the circuit produced by the Sorted Edges algorithm using the graph below.



Answer

AB: Add, cost 11

BG: Add, cost 13

AE: Add, cost 14

EF: Add, cost 15

EC: Skip (degree 3 at E)

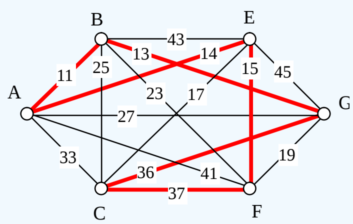
FG: Skip (would create a circuit not including C)

BF, BC, AG, AC: Skip (would cause a vertex to have degree 3)

GC: Add, cost 36

CF: Add, cost 37, completes the circuit

Final circuit: ABGCFEA



[1] There are some theorems that can be used in specific circumstances, such as Dirac's theorem, which says that a Hamiltonian circuit must exist on a graph with *n* vertices if each vertex has degree *n*/2 or greater.

4.3: Shortest Path

When you visit a website like Google Maps or use your Smartphone to ask for directions from home to your Aunt's house in Pasadena, you are usually looking for a shortest path between the two locations. These computer applications use representations of the street maps as graphs, with estimated driving times as edge weights.

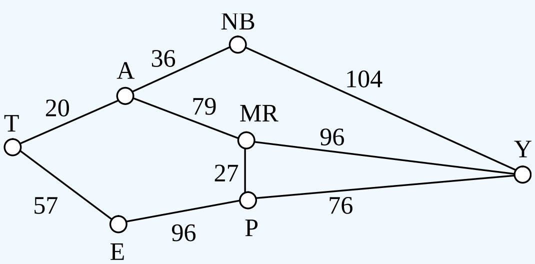
While often it is possible to find a shortest path on a small graph by guess-and-check, our goal in this chapter is to develop methods to solve complex problems in a systematic way by following **algorithms**. An algorithm is a step-by-step procedure for solving a problem. Dijkstra's (pronounced dike-stra) algorithm will find the shortest path between two vertices.

Dijkstra's Algorithm

1. Mark the ending vertex with a distance of zero. Designate this vertex as current.
2. Find all vertices leading to the current vertex. Calculate their distances to the end. Since we already know the distance the current vertex is from the end, this will just require adding the most recent edge. Don't record this distance if it is longer than a previously recorded distance.
3. Mark the current vertex as visited. We will never look at this vertex again.
4. Mark the vertex with the smallest distance as current, and repeat from step 2.

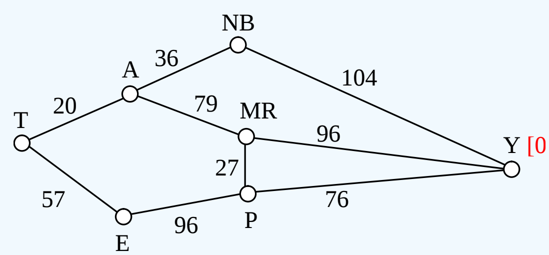
Example 3

Suppose you need to travel from Tacoma, WA (vertex T) to Yakima, WA (vertex Y). Looking at a map, it looks like driving through Auburn (A) then Mount Rainier (MR) might be shortest, but it's not totally clear since that road is probably slower than taking the major highway through North Bend (NB). A graph with travel times in minutes is shown below. An alternate route through Eatonville (E) and Packwood (P) is also shown.

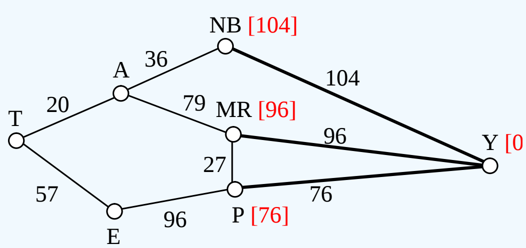


Solution

Step 1: Mark the ending vertex with a distance of zero. The distances will be recorded in [brackets] after the vertex name.



Step 2: For each vertex leading to Y, we calculate the distance to the end. For example, NB is a distance of 104 from the end, and MR is 96 from the end. Remember that distances in this case refer to the travel time in minutes.

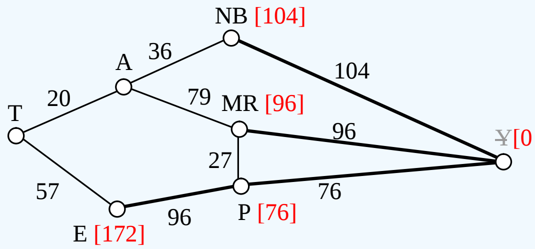


Step 3 & 4: We mark Y as visited, and mark the vertex with the smallest recorded distance as current. At this point, P will be designated current. Back to step 2.

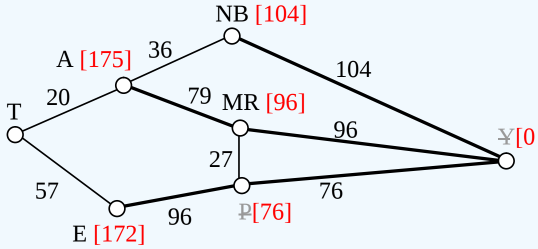
Step 2 (#2): For each vertex leading to P (and not leading to a visited vertex) we find the distance from the end. Since E is 96 minutes from P, and we've already calculated P is 76 minutes from Y, we can compute that E is 96+76 = 172 minutes from Y.

If we make the same computation for MR, we'd calculate 76+27 = 103. Since this is larger than the previously recorded distance from Y to MR, we will *not* replace it.

Step 3 & 4 (#2): We mark P as visited, and designate the vertex with the smallest recorded distance as current: MR. Back to step 2.

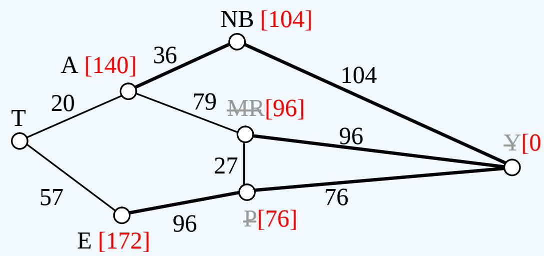


Step 2 (#3): For each vertex leading to MR (and not leading to a visited vertex) we find the distance to the end. The only vertex to be considered is A, since we've already visited Y and P. Adding MR's distance 96 to the length from A to MR gives the distance 96+79 = 175 minutes from A to Y.



Step 3 & 4 (#3): We mark MR as visited, and designate the vertex with smallest recorded distance as current: NB. Back to step 2.

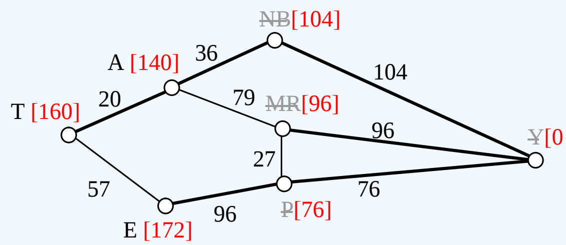
Step 2 (#4): For each vertex leading to NB, we find the distance to the end. We know the shortest distance from NB to Y is 104 and the distance from A to NB is 36, so the distance from A to Y through NB is 104+36 = 140. Since this distance *is* shorter than the previously calculated distance from Y to A through MR, we replace it.



Step 3 & 4 (#4): We mark NB as visited, and designate A as current, since it now has the shortest distance.

Step 2 (#5): T is the only non-visited vertex leading to A, so we calculate the distance from T to Y through A: 20+140 = 160 minutes.

Step 3 & 4 (#5): We mark A as visited, and designate E as current.



Step 2 (#6): The only non-visited vertex leading to E is T. Calculating the distance from T to Y through E, we compute $172 + 57 = 229$ minutes. Since this is longer than the existing marked time, we do not replace it.

Step 3 (#6): We mark E as visited. Since all vertices have been visited, we are done.

From this, we know that the shortest path from Tacoma to Yakima will take 160 minutes. Tracking which sequence of edges yielded 160 minutes, we see the shortest path is T-A-NB-Y.

Dijkstra's algorithm is an **optimal algorithm**, meaning that it always produces the actual shortest path, not just a path that is pretty short, provided one exists. This algorithm is also **efficient**, meaning that it can be implemented in a reasonable amount of time. Dijkstra's algorithm takes around V^2 calculations, where V is the number of vertices in a graph[1]. A graph with 100 vertices would take around 10,000 calculations. While that would be a lot to do by hand, it is not a lot for computer to handle. It is because of this efficiency that your car's GPS unit can compute driving directions in only a few seconds.

In contrast, an **inefficient** algorithm might try to list all possible paths then compute the length of each path. Trying to list all possible paths could easily take 10^{25} calculations to compute the shortest path with only 25 vertices; that's a 1 with 25 zeros after it! To put that in perspective, the fastest computer in the world would still spend over 1000 years analyzing all those paths.

✔ Example 4

A shipping company needs to route a package from Washington, D.C. to San Diego, CA. To minimize costs, the package will first be sent to their processing center in Baltimore, MD then sent as part of mass shipments between their various processing centers, ending up in their processing center in Bakersfield, CA. From there it will be delivered in a small truck to San Diego.

The travel times, in hours, between their processing centers are shown in the table below. Three hours has been added to each travel time for processing. Find the shortest path from Baltimore to Bakersfield.

	Baltimore	Denver	Dallas	Chicago	Atlanta	Bakersfield
Baltimore	*			15	14	
Denver		*		18	24	19
Dallas			*	18	15	25
Chicago	15	18	18	*	14	
Atlanta	14	24	15	14	*	
Bakersfield		19	25			*

Solution

While we could draw a graph, we can also work directly from the table.

Step 1: The ending vertex, Bakersfield, is marked as current.

Step 2: All cities connected to Bakersfield, in this case Denver and Dallas, have their distances calculated; we'll mark those distances in the column headers.

Step 3 & 4: Mark Bakersfield as visited. Here, we are doing it by shading the corresponding row and column of the table. We mark Denver as current, shown in bold, since it is the vertex with the shortest distance.

	Baltimore	**Denver** [19]	Dallas [25]	Chicago	Atlanta	Bakersfield [0]
Baltimore	*			15	14	
Denver		*		18	24	19
Dallas			*	18	15	25
Chicago	15	18	18	*	14	
Atlanta	14	24	15	14	*	
Bakersfield		19	25			*

Step 2 (#2): For cities connected to Denver, calculate distance to the end. For example, Chicago is 18 hours from Denver, and Denver is 19 hours from the end, the distance for Chicago to the end is 18+19 = 37 (Chicago to Denver to Bakersfield). Atlanta is 24 hours from Denver, so the distance to the end is 24+19 = 43 (Atlanta to Denver to Bakersfield).

Step 3 & 4 (#2): We mark Denver as visited and mark Dallas as current.

	Baltimore	Denver [19]	**Dallas** [25]	Chicago [37]	Atlanta [43]	Bakersfield [0]
Baltimore	*			15	14	
Denver		*		18	24	19
Dallas			*	18	15	25
Chicago	15	18	18	*	14	
Atlanta	14	24	15	14	*	
Bakersfield		19	25			*

Step 2 (#3): For cities connected to Dallas, calculate the distance to the end. For Chicago, the distance from Chicago to Dallas is 18 and from Dallas to the end is 25, so the distance from Chicago to the end through Dallas would be 18+25 = 43. Since this is longer than the currently marked distance for Chicago, we do not replace it. For Atlanta, we calculate 15+25 = 40. Since this is shorter than the currently marked distance for Atlanta, we replace the existing distance.

Step 3 & 4 (#3): We mark Dallas as visited, and mark Chicago as current.

	Baltimore	Denver [19]	Dallas [25]	**Chicago** [37]	Atlanta [40]	Bakersfield [0]
Baltimore	*			15	14	
Denver		*		18	24	19
Dallas			*	18	15	25
Chicago	15	18	18	*	14	
Atlanta	14	24	15	14	*	
Bakersfield		19	25			*

Step 2 (#4): Baltimore and Atlanta are the only non-visited cities connected to Chicago. For Baltimore, we calculate 15+37 = 52 and mark that distance. For Atlanta, we calculate 14+37 = 51. Since this is longer than the existing distance of 40 for Atlanta, we do not replace that distance.

Step 3 & 4 (#4): Mark Chicago as visited and Atlanta as current.

	Baltimore [52]	Denver [19]	Dallas [25]	Chicago [37]	**Atlanta** [40]	Bakersfield [0]
Baltimore	*			15	14	
Denver		*		18	24	19
Dallas			*	18	15	25
Chicago	15	18	18	*	14	
Atlanta	14	24	15	14	*	
Bakersfield		19	25			*

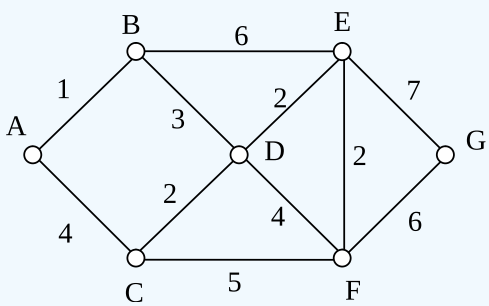
Step 2 (#5): The distance from Atlanta to Baltimore is 14. Adding that to the distance already calculated for Atlanta gives a total distance of 14+40 = 54 hours from Baltimore to Bakersfield through Atlanta. Since this is larger than the currently calculated distance, we do not replace the distance for Baltimore.

Step 3 & 4 (#5): We mark Atlanta as visited. All cities have been visited and we are done.

The shortest route from Baltimore to Bakersfield will take 52 hours, and will route through Chicago and Denver.

? Try it Now 2

Find the shortest path between vertices A and G in the graph below.



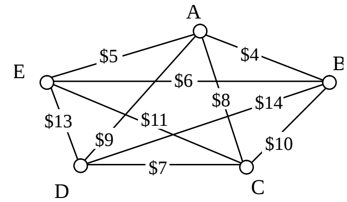
Answer

The shortest path is ABDEG, with length 13

[1] It can be made to run faster through various optimizations to the implementation.

4.4: Spanning Trees

A company requires reliable internet and phone connectivity between their five offices (named A, B, C, D, and E for simplicity) in New York, so they decide to lease dedicated lines from the phone company. The phone company will charge for each link made. The costs, in thousands of dollars per year, are shown in the graph.



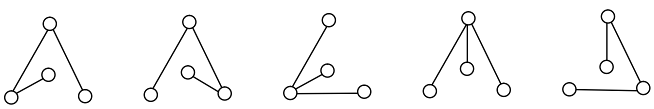
In this case, we don't need to find a circuit, or even a specific path; all we need to do is make sure we can make a call from any office to any other. In other words, we need to be sure there is a path from any vertex to any other vertex.

Spanning Tree

A **spanning tree** is a connected graph using all vertices in which there are no circuits.

In other words, there is a path from any vertex to any other vertex, but no circuits.

Some examples of spanning trees are shown below. Notice there are no circuits in the trees, and it is fine to have vertices with degree higher than two.



Usually we have a starting graph to work from, like in the phone example above. In this case, we form our spanning tree by finding a **subgraph** – a new graph formed using all the vertices but only some of the edges from the original graph. No edges will be created where they didn't already exist.

Of course, any random spanning tree isn't really what we want. We want the **minimum cost spanning tree (MCST)**.

Minimum Cost Spanning Tree (MCST)

The minimum cost spanning tree is the spanning tree with the smallest total edge weight.

A nearest neighbor style approach doesn't make as much sense here since we don't need a circuit, so instead we will take an approach similar to sorted edges.

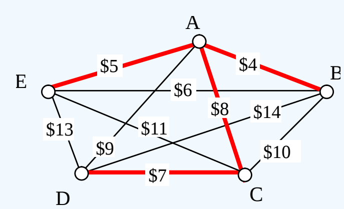
Kruskal's Algorithm

1. Select the cheapest unused edge in the graph.
2. Repeat step 1, adding the cheapest unused edge, unless:
   a. adding the edge would create a circuit
3. Repeat until a spanning tree is formed

Example 22

Using our phone line graph from above, begin adding edges:

AB	$4	OK
AE	$5	OK
BE	$6	reject-closes circuit ABEA
DC	$7	OK
AC	$8	OK



At this point we stop – every vertex is now connected, so we have formed a spanning tree with cost $24 thousand a year.

Remarkably, Kruskal's algorithm is both optimal and efficient; we are guaranteed to always produce the optimal MCST.

Example 23

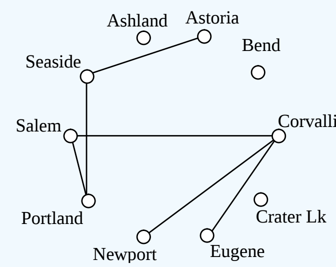
The power company needs to lay updated distribution lines connecting the ten Oregon cities below to the power grid. How can they minimize the amount of new line to lay?

	Ashland	Astoria	Bend	Corvallis	Crater Lake	Eugene	Newport	Portland	Salem	Seaside
Ashland	_	374	200	223	108	178	252	285	240	356
Astoria	374	_	255	166	433	199	135	95	136	17
Bend	200	255	_	128	277	128	180	160	131	247
Corvallis	223	166	128	_	430	47	52	84	40	155
Crater Lake	108	433	277	430	_	453	478	344	389	423
Eugene	178	199	128	47	453	_	91	110	64	181
Newport	252	135	180	52	478	91	_	114	83	117
Portland	285	95	160	84	344	110	114	_	47	78
Salem	240	136	131	40	389	64	83	47	_	118
Seaside	356	17	247	155	423	181	117	78	118	_

Solution

Using Kruskal's algorithm, we add edges from cheapest to most expensive, rejecting any that close a circuit. We stop when the graph is connected.

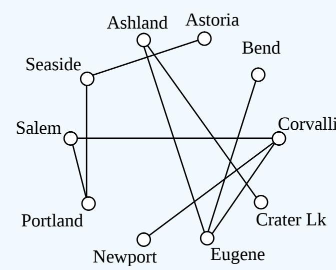
Seaside to Astoria	17 miles
Corvallis to Salem	40 miles
Portland to Salem	47 miles
Corvallis to Eugene	47 miles
Corvallis to Newport	52 miles
Salem to Eugene	reject – closes circuit
Portland to Seaside	78 miles



The graph up to this point is shown to the right.

Continuing,

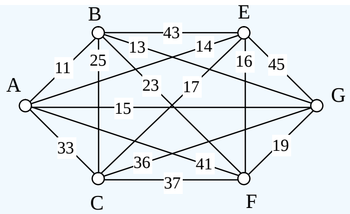
Newport to Salem	reject
Corvallis to Portland	reject
Eugene to Newport	reject
Portland to Astoria	reject
Ashland to Crater Lk	108 miles
Eugene to Portland	reject
Newport to Portland	reject
Newport to Seaside	reject
Salem to Seaside	reject
Bend to Eugene	128 miles
Bend to Salem	reject
Astoria to Newport	reject
Salem to Astoria	reject
Corvallis to Seaside	reject
Portland to Bend	reject
Astoria to Corvallis	reject
Eugene to Ashland	178 miles



This connects the graph. The total length of cable to lay would be 695 miles.

? Try it Now 7

Find a minimum cost spanning tree on the graph below using Kruskal's algorithm.



Answer

AB: Add, cost 11

BG: Add, cost 13

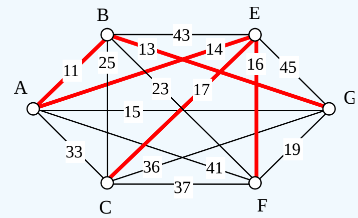
AE: Add, cost 14

AG: Skip, would create circuit ABGA

EF: Add, cost 16

EC: Add, cost 17

This completes the spanning tree



CHAPTER OVERVIEW

5: Numeration Systems

In the following chapter, we will try to focus on two main ideas. The first will be an examination of basic number and counting systems and the symbols that we use for numbers. We will look at our own modern (Western) number system as well those of a couple of selected civilizations to see the differences and diversity that is possible when humans start counting. The second idea we will look at will be base systems. By comparing our own base-ten (decimal) system with other bases, we will quickly become aware that the system that we are so used to, when slightly changed, will challenge our notions about numbers and what symbols for those numbers actually mean.

Thumbnail: Roman Numerals. (Public Domain; Monaneko via Wikipedia)

5.1: The Evolution of Numeration Systems

Introduction

As we begin our journey through the history of mathematics, one question to be asked is "Where do we start?" Depending on how you view mathematics or numbers, you could choose any of a number of launching points from which to begin. Howard Eves suggests the following list of possibilities.[i]

Where to start the study of the history of mathematics…

- At the first logical geometric "proofs" traditionally credited to Thales of Miletus (600 BCE).
- With the formulation of methods of measurement made by the Egyptians and Mesopotamians/Babylonians.
- Where prehistoric peoples made efforts to organize the concepts of size, shape, and number.
- In pre-human times in the very simple number sense and pattern recognition that can be displayed by certain animals, birds, etc.
- Even before that in the amazing relationships of numbers and shapes found in plants.
- With the spiral nebulae, the natural course of planets, and other universe phenomena.

We can choose no starting point at all and instead agree that mathematics has *always* existed and has simply been waiting in the wings for humans to discover. Each of these positions can be defended to some degree and which one you adopt (if any) largely depends on your philosophical ideas about mathematics and numbers.

Nevertheless, we need a starting point. Without passing judgment on the validity of any of these particular possibilities, we will choose as our starting point the emergence of the idea of number and the process of counting as our launching pad. This is done primarily as a practical matter given the nature of this course. In the following chapter, we will try to focus on two main ideas. The first will be an examination of basic number and counting systems and the symbols that we use for numbers. We will look at our own modern (Western) number system as well those of a couple of selected civilizations to see the differences and diversity that is possible when humans start counting. The second idea we will look at will be base systems. By comparing our own base-ten (decimal) system with other bases, we will quickly become aware that the system that we are so used to, when slightly changed, will challenge our notions about numbers and what symbols for those numbers actually mean.

Recognition of More vs. Less

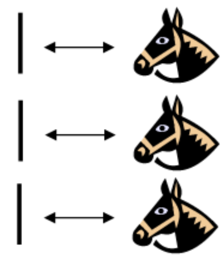
The idea of numbers and the process of counting goes back far beyond when history began to be recorded. There is some archeological evidence that suggests that humans were counting as far back as 50,000 years ago.[ii] However, we do not really know how this process started or developed over time. The best we can do is to make a good guess as to how things progressed. It is probably not hard to believe that even the earliest humans had some sense of *more* and *less*. Even some small animals have been shown to have such a sense. For example, one naturalist tells of how he would secretly remove one egg each day from a plover's nest. The mother was diligent in laying an extra egg every day to make up for the missing egg. Some research has shown that hens can be trained to distinguish between even and odd numbers of pieces of food.[iii] With these sorts of findings in mind, it is not hard to conceive that early humans had (at least) a similar sense of more and less. However, our conjectures about how and when these ideas emerged among humans are simply that; educated guesses based on our own assumptions of what might or could have been.

The Need for Simple Counting

As societies and humankind evolved, simply having a sense of more or less, even or odd, etc., would prove to be insufficient to meet the needs of everyday living. As tribes and groups formed, it became important to be able to know how many members were in the group, and perhaps how many were in the enemy's camp. Certainly it was important for them to know if the flock of sheep or other possessed animals were increasing or decreasing in size. "Just how many of them do we have, anyway?" is a question that we do not have a hard time imagining them asking themselves (or each other).

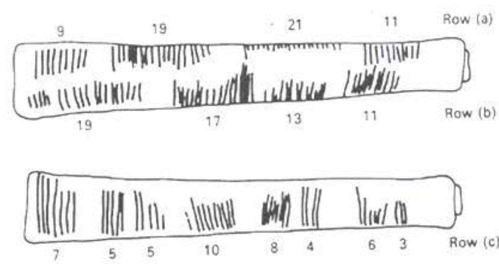
In order to count items such as animals, it is often conjectured that one of the earliest methods of doing so would be with "tally sticks." These are objects used to track the numbers of items to be counted. With this method, each "stick" (or pebble, or whatever counting device being used) represents one animal or object.

In the picture to the right, you see each stick corresponding to one horse. By examining the collection of sticks in hand one knows how many animals should be present. You can imagine the usefulness of such a system, at least for smaller numbers of items to keep track of. If a herder wanted to “count off” his animals to make sure they were all present, he could mentally (or methodically) assign each stick to one animal and continue to do so until he was satisfied that all were accounted for.



Of course, in our modern system, we have replaced the sticks with more abstract objects. In particular, the top stick is replaced with our symbol “1,” the second stick gets replaced by a “2” and the third stick is represented by the symbol “3,” but we are getting ahead of ourselves here. These modern symbols took many centuries to emerge.

Another possible way of employing the “tally stick” counting method is by making marks or cutting notches into pieces of wood, or even tying knots in string (as we shall see later). In 1937, Karl Absolom discovered a wolf bone that goes back possibly 30,000 years. It is believed to be a counting device.[iv] Another example of this kind of tool is the Ishango Bone, discovered in 1960 at Ishango, and shown below.[v] It is reported to be between six and nine thousand years old and shows what appear to be markings used to do counting of some sort.



The markings on rows (a) and (b) each add up to 60. Row (b) contains the prime numbers between 10 and 20. Row (c) seems to illustrate for the method of doubling and multiplication used by the Egyptians. It is believed that this may also represent a lunar phase counter.

Spoken Words

As methods for counting developed, and as language progressed as well, it is natural to expect that spoken words for numbers would appear. Unfortunately, the developments of these words, especially those corresponding to the numbers from one through ten, are not easy to trace. Past ten, however, we do see some patterns:

Eleven comes from “ein lifon,” meaning “one left over.”

Twelve comes from “twe lif,” meaning “two left over.”

Thirteen comes from “Three and ten” as do fourteen through nineteen.

Twenty appears to come from “twe-tig” which means “two tens.”

Hundred probably comes from a term meaning “ten times.”

Written Numbers

When we speak of “written” numbers, we have to be careful because this could mean a variety of things. It is important to keep in mind that modern paper is only a little more than 100 years old, so “writing” in times past often took on forms that might look quite unfamiliar to us today.

As we saw earlier, some might consider wooden sticks with notches carved in them as writing as these are means of recording information on a medium that can be “read” by others. Of course, the symbols used (simple notches) certainly did not leave a lot of flexibility for communicating a wide variety of ideas or information.

Other mediums on which “writing” may have taken place include carvings in stone or clay tablets, rag paper made by hand (12^{th} century in Europe, but earlier in China), papyrus (invented by the Egyptians and used up until the Greeks), and parchments from animal skins. And these are just a few of the many possibilities.

These are just a few examples of early methods of counting and simple symbols for representing numbers. Extensive books, articles and research have been done on this topic and could provide enough information to fill this entire course if we allowed it to. The range and diversity of creative thought that has been used in the past to describe numbers and to count objects and people is staggering. Unfortunately, we don’t have time to examine them all, but it is fun and interesting to look at some systems in more detail.

Egyptian Hieroglyphs

The system of ancient Egyptian numerals was used in Ancient Egypt from around 3000 BCE until the early first millennium CE. It was a system of numeration based on multiples of ten written in hieroglyphs. The Egyptians had no concept of a place-valued system such as the decimal system. The hieratic form of numerals stressed an exact finite series notation, ciphered one-to-one onto the Egyptian alphabet.

These symbols represent the powers of ten. From left to right: 1 (stroke), 10 (heel bone), 100 (scroll), 1000 (lotus flower), 10,000 (finger), 100,000 (tadpole), 1,000,000 (astonished person).

They used a simple grouping system by combining multiple symbols of the same value to add them up.

Example 1

Convert to our decimal system:

Solution

Since there are 3 scrolls, 4 heel bones, and 7 strokes, this represents the number 347.

The Egyptians also had a method for multiplication. This method is based on the fact that every positive integer can be written as the sum of different powers of 2.

Powers of 2

Powers of 2	Value
2^0	1
2^1	2
2^2	4
2^3	8
2^4	16
2^5	32
2^6	64

Example 2

Write the numbers 35 and 79 as a sum of powers of 2.

Solution

$35 = 2^5 + 2^1 + 2^0 = 32 + 2 + 1$

$79 = 2^6 + 2^3 + 2^2 + 2^1 + 2^0 = 64 + 8 + 4 + 2 + 1$

Example 3

Multiply 35×79

Solution

$35 \times 79 \ = 35 \times (64 + 8 + 4 + 2 + 1)$

$35 \times 79 \ = 35 \times 64 + 35 \times 8 + 35 \times 4 + 35 \times 2 + 35 \times 1$

Now make a table of Powers of 2 times 35 just by doubling 35 many times.

Powers of 2 Times 35

Powers of 2	Times 35
1	35
2	70
4	140
8	280
16	560
32	1120
64	2240

Thus $35 \times 79 \ = 2240 + 280 + 140 + 70 + 35 = 2765$.

Roman Numerals

Even though Roman Numerals are rare in today's society, they are still used and expected to be understood. They are taught in grades three through five, depending on the district. They can be seen in clocks, the Super Bowl, Film Credits for the copyright date like MCMLXII, preface of textbooks and others like Star Wars Episode VI and WWII.

Basic Table of Roman Numerals

Value	Symbol
1	I
5	V
10	X
50	L
100	C
500	D
1000	M

An line over a numeral means to multiply by 1000. You'll see in this next table.

The Complete Table of Roman Numerals

	1	2	3	4	5	6	7	8	9
ONES	I	II	III	IV	V	VI	VII	VIII	IX
TENS	X	XX	XXX	XL	L	LX	LXX	LXXX	XC
HUNDREDS	C	CC	CCC	CD	D	DC	DCC	DCCC	CM
THOUSANDS	M	MM	MMM	$\overline{\text{IV}}$	$\overline{\text{V}}$	$\overline{\text{VI}}$	$\overline{\text{VII}}$	$\overline{\text{VIII}}$	$\overline{\text{IX}}$
TEN THOUSANDS	$\bar{X}$	$\overline{XX}$	$\overline{XXX}$	$\overline{XL}$	$\bar{L}$	$\overline{LX}$	$\overline{LXX}$	$\overline{LXXX}$	$\overline{XC}$

	1	2	3	4	5	6	7	8	9
HUNDRED THOUSANDS	$\bar{C}$	$\overline{CC}$	$\overline{CCC}$	$\overline{CD}$	$\bar{D}$	$\overline{DC}$	$\overline{DCC}$	$\overline{DCCC}$	$\overline{CM}$

Example 4

For all numbers except 4 and 9, we **ADD** the Roman Numerals together, in order from left to right, greatest value to lowest value.

$$\begin{aligned} 11 &= 10+1 = XI \\ 8 &= 5+1+1+1 = VIII \\ 123 &= 100+10+10+1+1+1 = CXXIII \\ 3816 &= 1000+1000+1000+500+100+100+100+10+5+1 = MMMDCCCXVI \\ 7002 &= 5000+1000+1000+1+1 = \overline{V}\overline{II}II \end{aligned}$$

Example 5

For any number that includes a 4 or a 9, we subtract. When we are looking at a Roman Number expression and we see a Roman character **OUT OF ORDER**, which is the clue to **SUBTRACT**!

$$\begin{aligned} 4 &= 5-1 = IV \\ 9 &= 10-1 = IX \\ 40 &= 50-10 = XL \\ 99 &= (100-10)+(10-1) = XCIX \\ 400 &= 500-100 = CD \end{aligned}$$

Notice $99 \neq 100-1$ because you can only subtract numerals from the next two higher numerals.

---

[i] Eves, Howard; An Introduction to the History of Mathematics, p. 9.

[ii] Eves, p. 9.

[iii] McLeish, John; The Story of Numbers - How Mathematics Has Shaped Civilization, p. 7.

[iv] Bunt, Lucas; Jones, Phillip; Bedient, Jack; The Historical Roots of Elementary Mathematics, p. 2.

[v] http://www.math.buffalo.edu/mad/Ancient-Africa/mad_zaire-uganda.html

5.2: Place Value Systems

The Positional System

Place Value System

In a **place value system,** also called a **positional system,** the placement of the symbols in a numeral determines the value of the symbols.

For example in the numeral 73, the 7 represents seven tens and the 3 represents three ones, or seventy-three. However, if we switched the order to 37, the 3 would represent three tens and the 7 would represent seven ones, or thirty-seven.

The number system we use today is the Hindu-Arabic system which uses a place value system involving powers of ten. (Base 10) This was not the first place value system and civilizations used bases other than 10. As you might imagine, the development of a base system is an important step in making the counting process more efficient. Our own base-ten system probably arose from the fact that we have 10 fingers (including thumbs) on two hands. This is a natural development. However, other civilizations have had a variety of bases other than ten. For example, the Natives of Queensland used a base-two system, counting as follows: “one, two, two and one, two two’s, much.” Some Modern South American Tribes have a base-five system counting in this way: “one, two, three, four, hand, hand and one, hand and two,” and so on. The Babylonians used a base-sixty (sexigesimal) system. We wrap up with a specific example of a civilization that actually used a base system other than 10.

The Mayan Civilization

1. Chichen Itza
2. Uxmal
3. Tulum
4. Palenque
5. Bonampak, Yaxchilan
6. Tikal
7. Altun Ha
8. Copán

The Mayan civilization is generally dated from 1500 B.C.E to 1700 C.E. The Yucatan Peninsula (see map[i]) in Mexico was the scene for the development of one of the most advanced civilizations of the ancient world. The Mayans had a sophisticated ritual system that was overseen by a priestly class. This class of priests developed a philosophy with time as divine and eternal.[ii] The calendar, and calculations related to it, were thus very important to the ritual life of the priestly class, and hence the Mayan people. In fact, much of what we know about this culture comes from their calendar records and astronomy data. Another important source of information on the Mayans is the writings of Father Diego de Landa, who went to Mexico as a missionary in 1549.

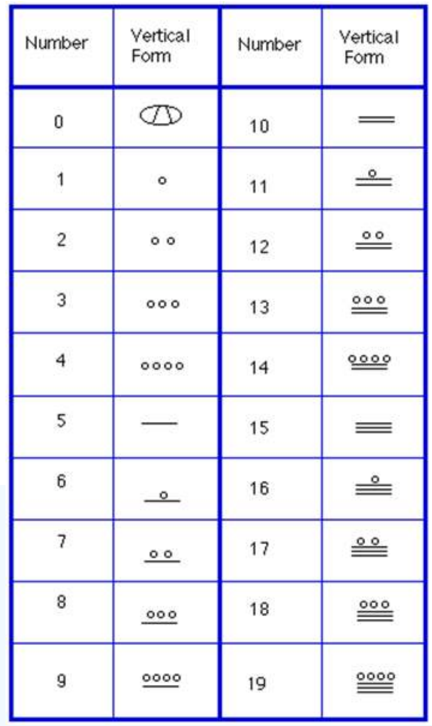
There were two numeral systems developed by the Mayans - one for the common people and one for the priests. Not only did these two systems use different symbols, they also used different base systems. For the priests, the number system was governed by ritual. The days of the year were thought to be gods, so the formal symbols for the days were decorated heads,[iii] like the sample to the left[iv] Since the basic calendar was based on 360 days, the priestly numeral system used a mixed base system employing multiples of 20 and 18, since $18 \times 20 = 360$.

The Mayan Number System

We will focus on the numeration system of the “priestly”. As we stated earlier, the Mayans used a slightly modified base-20 system, called the “vigesimal” system. Like our system, it is positional, meaning that the position of a numeric symbol indicates its place value. In the following table you can see the place value in its vertical format.[v]

Powers	Base-Ten Value
$20 \times 18 \times 20$	7200
20×18	360
20^1	20
20^0	1

In order to write numbers down, there were only three symbols needed in this system. A horizontal bar represented the quantity 5, a dot represented the quantity 1, and a special symbol (thought to be a shell) represented zero. The Mayan system may have been the first to make use of zero as a placeholder/number. The first 20 numbers are shown in the table to the right.[vi]

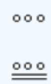


Unlike our system, where the ones place starts on the right and then moves to the left, the Mayan systems places the ones on the **bottom** of a vertical orientation and moves up as the place value increases.

When numbers are written in vertical form, there should never be more than four dots in a single place. When writing Mayan numbers, every group of five dots becomes one bar. Also, there should never be more than three bars in a single place…four bars would be converted to one dot in the next place up. It's the same as 10 getting converted to a 1 in the next place up when we carry during addition.

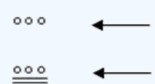
Example 1

What is the value of this number, which is shown in vertical form?



Solution

Starting from the bottom, we have the ones place. There are two bars and three dots in this place. Since each bar is worth 5, we have 13 ones when we count the three dots in the ones place. Looking to the place value above it (the twenties places), we see there are three dots so we have three twenties.



Hence we can write this number in base-ten as:

$$\begin{aligned}(3\times 20^1)+(13\times 20^0) &= (3\times 20)+(13\times 1)\\ &= 60+13\\ &= 73\end{aligned}$$

Example 2

What is the value of the following Mayan number?

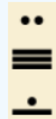
Solution

This number has 11 in the ones place, zero in the 20's place, and 18 in the 360's place. Hence, the value of this number in base-ten is:

$$18\times 360+0\times 20+11\times 1=6491$$

Try it Now 1

Convert the Mayan number below to base 10.



Answer

1026

Example 3

Convert the base 10 number 3757_{10} to Mayan numerals.

Solution

The highest place value that will divide into 3757 is 360 so we start by dividing that and then proceed from there:

$$\begin{aligned} 3757 \div 360 &= 10R157 \\ 157 \div 20 &= 7R17 \end{aligned}$$

This number indicates that we have 17 in the ones position. That's three bars and two dots at the bottom of the number. We also have 7 in the 20's place, so that's one bar and two dots in the second position. Finally, we have 10 in the 360's place, so that's two bars on the top. We get the following

Try it Now 2

Convert the base 10 number 10553_{10} to Mayan numerals.

Answer

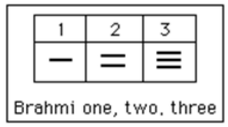
We have only scratched the surface of the wealth of research and information that exists on the development of numbers and counting throughout human history. What is important to note is that the system that we use every day is a product of thousands of years of progress and development. It represents contributions by many civilizations and cultures. It does not come down to us from the sky, a gift from the gods. It is not the creation of a textbook publisher. It is indeed as human as we are, as is the rest of mathematics. Behind every symbol, formula and rule there is a human face to be found, or at least sought.

The Evolution of the Hindu-Arabic System

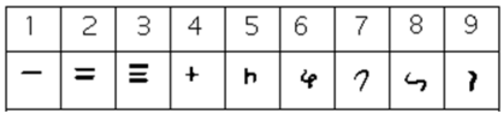
Our own number system, composed of the ten symbols {0,1,2,3,4,5,6,7,8,9} is called the *Hindu-Arabic system*. This is a base-ten (decimal) system since place values increase by powers of ten. Furthermore, this system is positional, which means that the position of a symbol has bearing on the value of that symbol within the number. For example, the position of the symbol 3 in the number 435,681 gives it a value much greater than the value of the symbol 8 in that same number. We'll explore base systems more thoroughly later. The development of these ten symbols and their use in a positional system comes to us primarily from India.[viii]

It was not until the 15th century that the symbols that we are familiar with today first took form in Europe. However, the history of these numbers and their development goes back hundreds of years. One important source of information on this topic is the writer al-Biruni, whose picture is shown here.[ix] Al-Biruni, who was born in modern day Uzbekistan, had visited India on several occasions and made comments on the Indian number system. When we look at the origins of the numbers that al-Biruni encountered, we have to go back to the third century B.C.E. to explore their origins. It is then that the Brahmi numerals were being used.

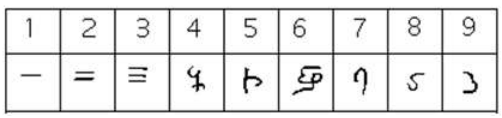
The Brahmi numerals were more complicated than those used in our own modern system. They had separate symbols for the numbers 1 through 9, as well as distinct symbols for 10, 100, 1000,…, also for 20, 30, 40,…, and others for 200, 300, 400, …, 900. The Brahmi symbols for 1, 2, and 3 are shown below.[x]



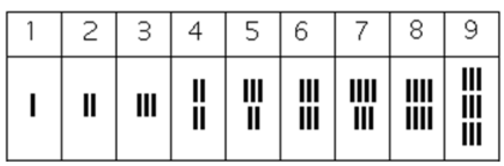
These numerals were used all the way up to the 4th century C.E., with variations through time and geographic location. For example, in the first century C.E., one particular set of Brahmi numerals took on the following form[xi]:



From the 4th century on, you can actually trace several different paths that the Brahmi numerals took to get to different points and incarnations. One of those paths led to our current numeral system, and went through what are called the Gupta numerals. The Gupta numerals were prominent during a time ruled by the Gupta dynasty and were spread throughout that empire as they conquered lands during the 4th through 6th centuries. They have the following form[xii]:

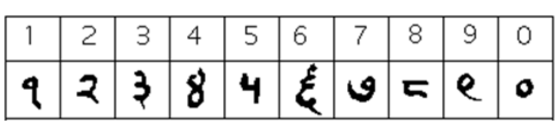


How the numbers got to their Gupta form is open to considerable debate. Many possible hypotheses have been offered, most of which boil down to two basic types[xiii]. The first type of hypothesis states that the numerals came from the initial letters of the names of the numbers. This is not uncommon…the Greek numerals developed in this manner. The second type of hypothesis states that they were derived from some earlier number system. However, there are other hypotheses that are offered, one of which is by the researcher Ifrah. His theory is that there were originally nine numerals, each represented by a corresponding number of vertical lines. One possibility is this:[xiv]



Because these symbols would have taken a lot of time to write, they eventually evolved into cursive symbols that could be written more quickly. If we compare these to the Gupta numerals above, we can try to see how that evolutionary process might have taken place, but our imagination would be just about all we would have to depend upon since we do not know exactly how the process unfolded.

The Gupta numerals eventually evolved into another form of numerals called the Nagari numerals, and these continued to evolve until the 11th century, at which time they looked like this:[xv]



Note that by this time, the symbol for 0 has appeared! The Mayans in the Americas had a symbol for zero long before this, however, as we shall see later in the chapter.

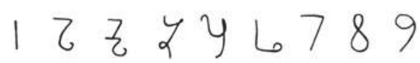
These numerals were adopted by the Arabs, most likely in the eighth century during Islamic incursions into the northern part of India.[xvi] It is believed that the Arabs were instrumental in spreading them to other parts of the world, including Spain (see below).

Other examples of variations up to the eleventh century include:

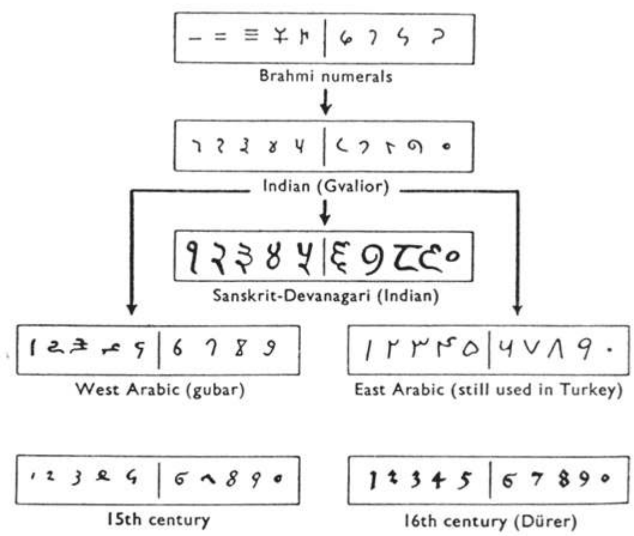
Devangari, eighth century[xvii]:

West Arab Gobar, tenth century[xviii]:

Spain, 976 B.C.E.[xix]:



Finally, one more graphic[xx] shows various forms of these numerals as they developed and eventually converged to the 15th century in Europe.

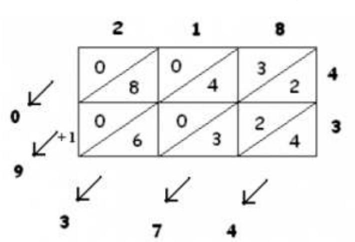


Galley Multiplication

In the 1500s in England, students were taught to compute multiplication following galley method.

To multiply 43 and 218, for example, draw a 2 × 3 grid of squares. Write the digits of the first number along the right side of the grid and the digits of the second number along the top.

Divide each cell of the grid diagonally and write in the product of the column digit and row digit of that cell, separating the tens from the units across the diagonal of that cell. (If the product is a one digit answer, place a 0 in the tens place.)

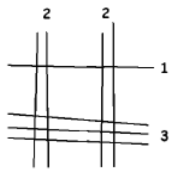


To get the answer, add the entries in each diagonal, carrying tens digits over to the next diagonal if necessary. In our example, we have

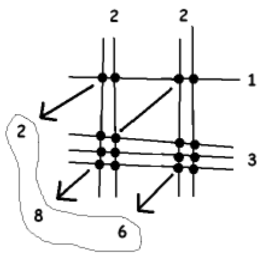
$218 \times 43 = 9374$.

Lines and Intersections

Here's an unusual way to perform multiplication. To compute 22 × 13, for example, draw two sets of vertical lines, the left set containing two lines and the right set two lines (for the digits in 22) and two sets of horizontal lines, the upper set containing one line and the lower set three (for the digits in 13).

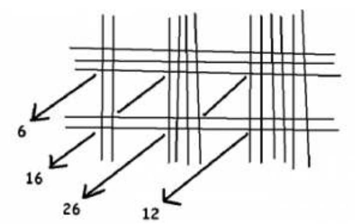


There are four sets of intersection points. Count the number of intersections in each and add the results diagonally as shown:



The answer 286 appears!

There is one possible glitch as illustrated by the computation 246 × 32:



Although the answer 6 thousands, 16 hundreds, 26 tens, and 12 ones is absolutely correct, one needs to carry digits and translate this as 7,872.

---

[i] www.gorp.com/gorp/location/latamer/map_maya.htm

[ii] Bidwell, James; Mayan Arithmetic in *Mathematics Teacher*, Issue 74 (Nov., 1967), p. 762-68.

[iii] www.ukans.edu/~lctls/Mayan/numbers.html

[iv] www.ukans.edu/~lctls/Mayan/numbers.html

[v] Bidwell

[vi] www.vpds.wsu.edu/fair_95/gym/UM001.html

[vii] forum.swarthmore.edu/k12/mayan.math/mayan2.html

[viii] www-groups.dcs.st-and.ac.uk/~history/HistTopics/Indian_numerals.html

[ix] www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Al-Biruni.html

[x] www-groups.dcs.st-and.ac.uk/~history/HistTopics/Indian_numerals.html

[xi] www-groups.dcs.st-and.ac.uk/~history/HistTopics/Indian_numerals.html

[xii] Ibid

[xiii] Ibid

[xiv] Ibid

[xv] Ibid

[xvi] Katz, page 230

[xvii] Burton, David M., *History of Mathematics, An Introduction*, p. 254-255

[xviii] Ibid

[xix] Ibid

[xx] Katz, page 231.

5.3: Calculating in Other Bases

Introduction and Basics

During the previous discussions, we have been referring to positional base systems. In this section of the chapter, we will explore exactly what a base system is and what it means if a system is “positional.” We will do so by first looking at our own familiar, base-ten system and then deepen our exploration by looking at other possible base systems. In the next part of this section, we will journey back to Mayan civilization and look at their unique base system, which is based on the number 20 rather than the number 10.

A base system is a structure within which we count. The easiest way to describe a base system is to think about our own base-ten system. The base-ten system, which we call the “decimal” system, requires a total of ten different symbols/digits to write any number. They are, of course, 0, 1, 2, ….. 9.

The decimal system is also an example of a *positional* base system, which simply means that the position of a digit gives its place value. Not all civilizations had a positional system even though they did have a base with which they worked.

Below is the number line for **Base 10**. Notice how it is broken up into rows. When you get to the last digit - in this case 9 - you have to add one to the next place value to the left.

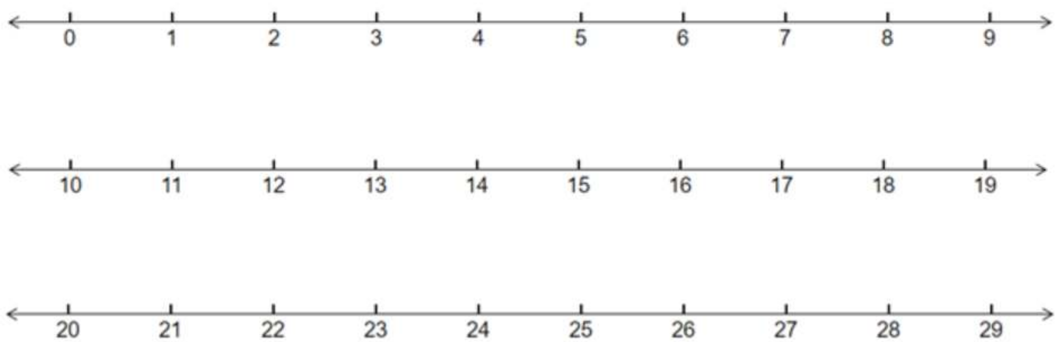


Figure 5.3.1: Number line for Base 10

There are an infinite amount of different bases and an infinite amount of corresponding number lines. Below are three different examples, written like we have Base 10 above.

Base 2 Number Line (reads from left to right, then top to bottom). The fifth number in the number line is 101_{two}.

Table 5.3.1: Base 2 Number Line

0	1
10	11
100	101
110	111
1000	1001
1010	1011
1100	1101
1110	1111
10000	10001

Base 8 Number Line (reads from left to right, then top to bottom). The 10^{th} number is 12_{eight}.

Table 5.3.2: Base 8 Number Line

0	1	2	3	4	5	6	7
10	11	12	13	14	15	16	17
20	21	22	23	24	25	26	27

Base 12 Number Line (reads from left to right, then top to bottom). The 15^{th} number is 13_{twelve}.

Table 5.3.3: Base 12 Number Line

0	1	2	3	4	5	6	7	8	9	A	B
10	11	12	13	14	15	16	17	18	19	1A	1B
20	21	22	23	24	25	26	27	28	29	2A	2B

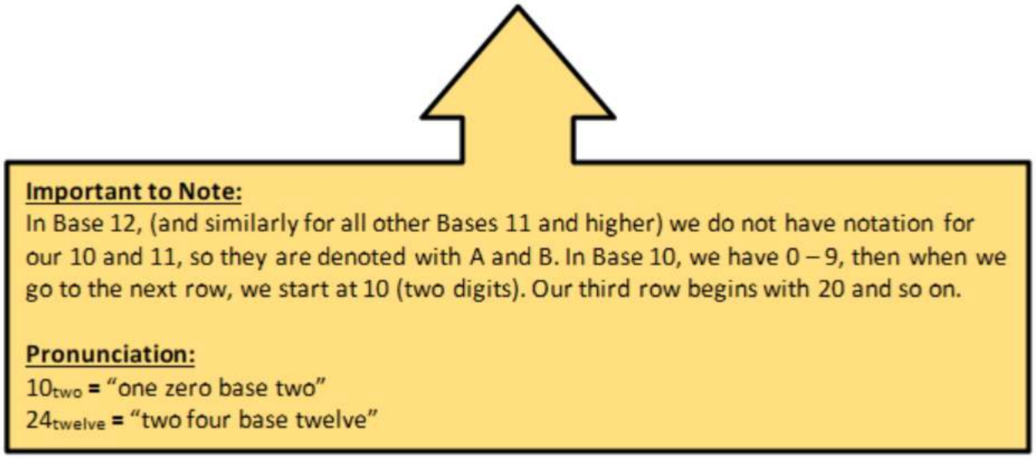


Figure 5.3.2

Where are different bases used?

Base 2: (Binary) Computers use Base 2, just zeros and ones for all of their programming. Your phone operates in only zeros (OFF) and ones (ON).

Base 5: (Quinary) This was one of the very first systems of counting, since we have five fingers.

Base 8: (Octal) Think of Base 8 as the mathematics for the Cartoon Universe. Many Cartoons have only eight fingers. We have 10 fingers and live in a Base 10 universe. Cartoon have eight fingers and live in a Base 8 universe. The Cartoon character does not have a character for eight or nine items, as we do in our universe.

Base 12: (Duodecimal) Base 12 is rare to find within history. However, we currently have 12 hours on the clock and 12 months in the year. There were a few tribes in Africa and India which used the duodecimal (base 12) system.

Base 16: (Hexadecimal) Base 16 is used in computers to represent long binary values. It allows you to store more information using less space.

Base 20: (Vigesimal) The Mayans used a Base 20 system, and invented the concept of zero.

Base 60: (Sexagesimal) An extreme example is base 60, which was used by the Babylonians (about 4000 years ago) which is now current day Iraq. They had characters for 1 – 59 items. (The concept of "zero" was not discovered yet.) However, this is where our current concept of 60 minutes and 60 seconds come from.

Converting From Other Bases To Base 10

In our base-ten system, a number like 5,783,216 has meaning to us because we are familiar with the system and its places. As we know, there are six ones, since there is a 6 in the ones place. Likewise, there are seven "hundred thousands," since the 7 resides in that place. Each digit has a value that is explicitly determined by its position within the number. We make a distinction between digit, which is just a symbol such as 5, and a number, which is made up of one or more digits. We can take this number and assign each of its digits a value. One way to do this is with a table, which follows:

$5,000,000$	$= 5 \times 1,000,000$	$= 5 \times 10^6$	Five million
$+700,000$	$= 7 \times 100,000$	$= 7 \times 10^5$	Seven hundred thousand
$+80,000$	$= 8 \times 10,000$	$= 8 \times 10^4$	Eighty thousand
$+3,000$	$= 3 \times 1000$	$= 3 \times 10^3$	Three thousand
$+200$	$= 2 \times 100$	$= 2 \times 10^2$	Two hundred
$+10$	$= 1 \times 10$	$= 1 \times 10^1$	Ten
$+6$	$= 6 \times 1$	$= 6 \times 10^0$	Six
$5,783,216$	Five million, seven hundred eighty-three thousand, two hundred sixteen		

From the third column in the table we can see that each place is simply a multiple of ten. Of course, this makes sense given that our base is ten. The digits that are multiplying each place simply tell us how many of that place we have. We are restricted to having at most 9 in any one place before we have to "carry" over to the next place. We cannot, for example, have 11 in the hundreds place. Instead, we would carry 1 to the thousands place and retain 1 in the hundreds place. This comes as no surprise to us since we readily see that 11 hundreds is the same as one thousand, one hundred. Carrying is a pretty typical occurrence in a base system.

However, base-ten is not the only option we have. Practically any positive integer greater than or equal to 2 can be used as a base for a number system. Such systems can work just like the decimal system except the number of symbols will be different and each position will depend on the base itself.

For example, let's suppose we adopt a base-five system. The only modern digits we would need for this system are 0,1,2,3 and 4. What are the place values in such a system? To answer that, we start with the ones place, as most base systems do. However, if we were to count in this system, we could only get to four (4) before we had to jump up to the next place. Our base is 5, after all! What is that next place that we would jump to? It would not be tens, since we are no longer in base-ten. We're in a different numerical world. As the base-ten system progresses from 10^0 to10^1, so the base-five system moves from 5^0 to $5^1 = 5$. Thus, we move from the ones to the fives.

After the fives, we would move to the 5^2 place, or the twenty fives. Note that in base-ten, we would have gone from the tens to the hundreds, which is, of course, 10^2.

Let's take an example and build a table. Consider the number 30412 in base five. We will write this as 30412_5 , where the subscript 5 is not part of the number but indicates the base we're using. First off, note that this is NOT the number "thirty thousand, four hundred twelve." We must be careful not to impose the base-ten system on this number. Here's what our table might look like. We will use it to convert this number to our more familiar base-ten system.

	Base 5	This column coverts to base-ten	In Base-Ten
	3×5^4	$= 3 \times 625$	$= 1875$
$+$	0×5^3	$= 0 \times 125$	$= 0$
$+$	4×5^2	$= 4 \times 25$	$= 100$
$+$	1×5^1	$= 1 \times 5$	$= 5$
$+$	2×5^0	$= 2 \times 1$	$= 2$
		Total	1982

As you can see, the number 30412_5 is equivalent to 1,982 in base-ten. We will say $30412_5 = 1982_{10}$. All of this may seem strange to you, but that's only because you are so used to the only system that you've ever seen.

Example 1

Convert 6234_7 to a base 10 number.

Solution

We first note that we are given a base-7 number that we are to convert. Thus, our places will start at the ones (7^0), and then move up to the $7's, 49's \left(7^2\right)$, etc. Here's the breakdown:

	Base 7	Convert	Base 10
	$= 6 \times 7^3$	$= 6 \times 343$	$= 2058$
$+$	$= 2 \times 7^2$	$= 2 \times 49$	$= 98$
$+$	$= 3 \times 7$	$= 3 \times 7$	$= 21$
$+$	$= 4 \times 1$	$= 4 \times 1$	$= 4$
		Total	2181

Thus $6234_7 = 2181_{10}$

Try it Now 1

Convert 41065_7 to a base 10 number.

Answer

$41065_7 = 9994_{10}$

Converting From Base 10 To Other Bases

Converting from an unfamiliar base to the familiar decimal system is not that difficult once you get the hang of it. It's only a matter of identifying each place and then multiplying each digit by the appropriate power. However, going the other direction can be a little trickier. Suppose you have a base-ten number and you want to convert to base-five. When converting from base-ten to some other base, it is helpful to determine the highest power of the base that will divide into the given number at least once.

Converting From Base 10 to Base b

1. Find the highest power of the base b that will divide into the given number at least once and then divide.
2. Write down the whole number part, then use the remainder from division in the next step.
3. Repeat step two, dividing by the next highest power of the base b, writing down the whole number part (including 0), and using the remainder in the next step.
4. Continue until the remainder is smaller than the base. This last remainder will be in the "ones" place.
5. Collect all your whole number parts to get your number in base b notation.

Example 2

Convert the base-ten number 348 to base-five.

Solution

The powers of five are:

$5^0 = 1$
$5^1 = 5$
$5^2 = 25$
$5^3 = 125$
$5^4 = 625$

Etc..

since 348 is smaller than 625, but bigger than 125, we see that $5^3 = 125$ is the highest power of five present in 348. So we divide 125 into 348 to see how many of them there are:

$348 \div 125 = 2$ with remainder 98

We write down the whole part, 2, and continue with the remainder. There are 98 left over, so we see how many 25's (the next smallest power of five) there are in the remainder:

$98 \div 25 = 3$ with remainder 23

We write down the whole part, 2, and continue with the remainder. There are 23 left over, so we look at the next place, the 5's:

$23 \div 5 = 4$ with remainder 3

This leaves us with 3, which is less than our base, so this number will be in the "ones" place. We are ready to assemble our base-five number:

$348 = \left(2 \times 5^3\right) + \left(3 \times 5^2\right) + \left(4 \times 5^1\right) + (3 \times 1)$

Hence, our base-five number is 2343. We'll say that $348_{10} = 2343_5$

Example 3

Convert the base-ten number 4509 to base-seven.

Solution

The powers of 7 are:

$7^0 = 1$
$7^1 = 7$
$7^2 = 49$
$7^3 = 343$
$7^4 = 2401$
$7^5 = 16807$

Etc…

The highest power of 7 that will divide into 4509 is $7^4 = 2401$

With division, we see that it will go in 1 time with a remainder of 2108. So we have 1 in the 7^4 place.

The next power down is $7^3 = 343$, which goes into 2108 six times with a new remainder of 50. So we have 6 in the 7^3 place.

The next power down is $7^2 = 49$, which goes into 50 once with a new remainder of 1. So there is a 1 in the 7^2 place.

The next power down is 7^1 but there was only a remainder of 1, so that means there is a 0 in the 7 's place and we still have 1 as a remainder.

That, of course, means that we have 1 in the ones place.

$$\begin{array}{llll} 4509 \div 7^4 = & 1 & \text{R} & 2108 \\ 2108 \div 7^3 = & 6 & \text{R} & 50 \\ 50 \div 7^2 = & 1 & \text{R} & 1 \\ 1 \div 7^1 = & 0 & \text{R} & 1 \\ 1 \div 7^0 = & 1 & & \\ 4509_{10} = 16101_7 & & & \end{array}$$

Putting all of this together means that $4509_{10} = 16101_7$

Try it Now 2

Convert 657_{10} to a base 4 number.

Answer

$657_{10} = 22101_4$

Try it Now 3

Convert 8377_{10} to a base 8 number.

Answer

$8377_{10} = 20271_8$

Operations In Other Bases

Adding In Other Bases

1. Rewrite the problem vertically (line up the place values).
2. Add the rightmost digits first (the *ones* place).
3. Represent the sum in the base you are working with and "carry" as if you are adding in Base 10.
4. Repeat steps 2 and 3 with the next place value higher until you get to the leftmost digit.

Example 4

$$5415_6 + 3042_6$$

$$\begin{array}{r} 5415_6 \\ +3042_6 \\ \hline 12501_6 \end{array}$$

To help us through this addition, here is a table of all possible sums of single-digits in base 6:

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	10
2	2	3	4	5	10	11
3	3	4	5	10	11	12
4	4	5	10	11	12	13
5	5	10	11	12	13	14

$5+2=7$ which is 11_6 so we "carry" the left 1 and put the right 1 in the *ones* place in the sum.

$1+1+4=6$ which is 10_6 so we "carry" the 1 and put the 0 in the *sixes* place in the sum.

$1+4+0=5$ which is 5_6 so we put the 5 in the *thirty-sixes* place in the sum.

$5+3=8$ which is 12_6 and it's the last step, so we put the 2 in the *two-hundred-sixteens* place and the 1 in the *one-thousand-two-hundred-ninety-sixes* place in the sum.

Example 5

$$1202_3 + 1022_3$$

$$\begin{array}{r} 1202_3 \\ +1022_3 \\ \hline 10001_3 \end{array}$$

$2+2=4$ which is 11_3 so we "carry" the left 1 and put the right 1 in the *ones* place.

$1+0+2=3$ which is 10_3 so we "carry" the 1 and put the 0 in the *threes* place.

$1+2+0=3$ which again is 10_3 so we again "carry" the 1 and put the 0 in the *nines* place.

$1+1+1=3$ which again is 10_3 and it's the last step, so we put the 0 in the *twenty-sevens* place and the 1 in the *eighty-ones* place.

Try it Now 4

Add $7452_9 + 3288_9$

Answer

11751_9

Subtraction In Other Bases

1. Rewrite the problem vertically (line up the place values).
2. Subtract the rightmost digits first (the *ones* place).
   - If you cannot get a positive value, you need to "borrow" from the next column over, exchanging 1 from the b column to get b in the ones column.
     - (For example in base 5: take 1 *five* and exchange it for 5 *ones*).
3. Repeat step 2 with the next place value higher until you get to the leftmost digit.

Example 6

$$4310_7 - 1234_7$$

$$\begin{array}{r} 4310_7 \\ -1234_7 \\ \hline 3043_7 \end{array}$$

We cannot take 4 from 0 so we take one from the *sevens* place and exchange it for 7 *ones*. $7-4=3$ so we put the 3 in the *ones* place in the difference.

We cannot take 3 from 0 (because we took 1) so we take one from the *forty-nines* place and exchange it for 7 *sevens*. $7-3=4$ so we put the 4 in the *sevens* place in the difference.

$2-2=0$ so we put the 0 in the *forty-nines* place in the difference.

$4-1=3$ so we put the 3 in the *three-hundred-forty-threes* place in the difference.

Example 7

$$3741_8 - 1465_8$$

$$\begin{array}{r} 3741_8 \\ -1465_8 \\ \hline 2254_8 \end{array}$$

We cannot take 5 from 1 so we take one from the *eights* place and exchange it for 8 ones and add it to the 1 already there. $9-5=4$ so we put the 3 in the *ones* place in the difference.

We cannot take 6 from 3 (because we took 1) so we take one from the *sixty-fours* place and exchange it for 8 *eights* and add it to the 3 already there. $11-6=5$ so we put the 5 in the *eights* place in the difference.

$6-4=2$ so we put the 2 in the *sixty-fours* place in the difference.

$3-1=2$ so we put the 2 in the *five-hundred-twelves* place in the difference.

Multiplication In Other Bases

1. Rewrite the problem vertically.
2. Multiply the bottom number's *ones* digit by the top number's *ones* digit.
3. Represent the product in the base you are working with and "carry" as if you are multiplying in Base 10.
4. Repeat steps 2 and 3 remembering to add the value of the "carry" until you have multiplies the bottom number's *ones* digit with every digit of the top number.
5. Now repeat steps 2 and 3 with the bottom number's next digit over. You need to place a 0 in the *ones* digit because the next digit is not a *ones* digit.
6. Repeat until you have use all digits of the bottom number.

Example 8

$$323_4 \times 12_4$$

$$\begin{array}{r} 323_4 \\ \times 12_4 \\ \hline 1312_4 \\ +3230_4 \\ \hline 11202_4 \end{array}$$

To help us through this multiplication, here is a table of all possible products of single-digits in base 4:

$\times$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	10	12
3	0	3	12	21

$2 \times 3 = 6$ which is 12_4 so we "carry" the 1 and put the 2 in the *ones* place in the product.

$2 \times 2 + 1 = 5$ which is 11_4 so we "carry" the left 1 and put the right 1 in the *fours* place in the product.

$2 \times 3 + 1 = 7$ which is 13_4 so we put the 3 in the *sixteens* place in the product and the 1 in the *sixty-fours* place in the product.

Now on the next line we put a 0 in *ones* place. Then we multiply by the 1 in 12_4.

$1 \times 3 = 3$ so we put the 3 in the *fours* place.

$1 \times 2 = 2$ so we put the 2 in the *sixteens* place.

$1 \times 3 = 3$ so we put the 3 in the *sixty-fours* place.

Now we add to get our final answer.

5.4: Modular Systems

Divides

Let a and b be two integers such that $a \neq 0$. We say "a **divides** b" or $a \mid b$ if and only if $b = aq$ for some integer q.

The following statements are equivalent:

- a ***divides*** b,
- a is a ***divisor*** of b,
- a is a ***factor*** of b,
- b is a ***multiple*** of a, and
- b is ***divisible by*** a.

In terms of division, we say that a divides b if and only if the remainder is zero when b is divided by a. We adopt the notation $a \mid b$.

Do not use a forward slash / or a backward slash \ in the notation. To say that a does not divide b, we add a slash across the vertical bar, as in $a \nmid b$

Do not confuse the notation $a \mid b$ with $\frac{a}{b}$. The notation $\frac{a}{b}$ represents a fraction. It is also written as a/b with a (forward) slash. It uses floating-point (that is, real or decimal) division. For example, $\frac{11}{4} = 2.75$.

Both integers a and b can be positive or negative, and b could even be 0. The only restriction is $a \neq 0$. In addition, q must be an integer.

Example 1

Check to see if these are true:

1. $5 \mid 35$
2. $7 \mid 14$
3. $8 \mid 36$

Solution

1. Yes $5 \mid 35$ because $35 = 5 \times 7$.
2. Yes $7 \mid 14$ because $14 = 7 \times 2$.
3. No $8 \nmid 36$ because $\frac{36}{8} = 4.5$ which is not an integer.

Modular arithmetic uses only a fixed number of possible results in all its computation. For instance, there are only 12 hours on the face of a clock. If the time now is 7 o'clock, 20 hours later will be 3 o'clock; and we do not say 27 o'clock! This example explains why modular arithmetic is referred to by some as ***clock arithmetic***.

Example 2

If the current time is 2:00 p.m. What time will it be 65 hours from now?

Solution

Assume the current time is 2:00 p.m. Write this as 14:00. Sixty five hours later, it would be 79:00. Since

$$79 = 24 \cdot 3 + 7,$$

it will be 7:00 or 7 a.m.

Example 3

If today is Wednesday. What day of the week will it be 11 days from now?

Solution

Designate Sunday, Monday, Tuesday, …, Saturday as Day 0, 1, 2, …, 6. If today is Wednesday, then today is Day 3. 11 days from now it will be day 14. Since

$$14 = 7 \cdot 2 + 0,$$

it will be Day 0 or Sunday.

In the clock example above, we essentially regard 27 o'clock the same as 3 o'clock. They key is, we are only interested in the remainder when a value is divided by 24. In the week example above, we are only interested in the remainder when a value is divided by 7.

Congruent Modulo m

Let $m \geq 2$ be a fixed integer. We say the two integers a and b are ***congruent modulo m***, denoted

$$a \equiv b \pmod{m}$$

if and only if $m \mid (a-b)$. The integer m is called the ***modulus*** of the congruence.

Example 4

Determine if these are true:

1. $59 \equiv 27 \pmod{8}$
2. $33 \equiv 11 \pmod{12}$
3. $11 \equiv 35 \pmod{6}$

Solution

1. Yes. $59-27=32$ and $8 \mid 32$.
2. No. $33-11=22$ and $\frac{22}{12} \approx 1.83$ which is not an integer.
3. Yes. $11-35=-24$ and $6 \mid -24$ because $-24=6 \times -4$.

Performing Arithmetic Operations In A Modulo m System

To add, subtract, and multiply in a modulo m system:

1. Perform the operation as usual.
2. Replace the result in (1) by one of the numbers 0, 1, 2, ... , $m-1$ that is congruent to that result.

(These are the remainders when dividing an integer by m).

Example 5

Perform the following operations:

1. $7+4 \pmod{8}$
2. $2-5 \pmod{12}$
3. $5 \times 8 \pmod{9}$

Solution

1. $7+4=11$ and $11 \equiv 3 \pmod{8}$, so $7+4 \equiv 3 \pmod{8}$.
2. $2-5=-3$ and $-3 \equiv 9 \pmod{12}$, ($-3+12=9$), so $2-5 \equiv 9 \pmod{12}$.
3. $5 \times 8=40$ and $40 \equiv 4 \pmod{9}$, ($40=9 \cdot 4+4$), so $5 \times 8 \equiv 4 \pmod{9}$.

✓ Example 6

Solve $4+x \equiv 2 \pmod{5}$ using trial and error.

Solution

We need to test all the remainders 0, 1, 2, 3, and 4 to see which one(s) work. Yes there can be multiple answers.

$4+0 \equiv 4 \pmod{5}$ FALSE

$4+1 \equiv 0 \pmod{5}$ FALSE

$4+2 \equiv 1 \pmod{5}$ FALSE

$4+3 \equiv 2 \pmod{5}$ TRUE

$4+4 \equiv 3 \pmod{5}$ FALSE

Thus $x \equiv 3 \pmod{5}$.

? Try it Now 1

Solve $2x \equiv 3 \pmod{6}$ by trial and error.

Answer

$x \equiv 0 \pmod{6}$ AND $x \equiv 3 \pmod{6}$.

✓ Example 7

In the late third century, the Chinese mathematician Sun Tzu asked his students: "We have things of which we do not know the number; if we count by threes, the remainder is 2; if we count be fives, the remainder is 3; if we count by sevens, the remainder is 2. How many things are there?"

Solution

Each of his sentences is a congruence. Let's set them up:

$$\begin{aligned} x &\equiv 2 \pmod{3} \\ x &\equiv 3 \pmod{5} \\ x &\equiv 2 \pmod{7} \end{aligned}$$

Then let's create a list of numbers that satisfy each congruence.

$$\begin{array}{r} 2,5,8,11,14,17,20,23,\ldots \\ 3,8,13,18,23,28,33,38,\ldots \\ 2,9,16,23,30,37,44,51,\ldots \end{array}$$

Since 23 is the smallest number that appears in all three lists, $x=23$ is the least positive number that satisfies all three congruences.

? Try it Now 2

Find the smallest positive integer such that if we divide by two, three, and four, the remainder is 1, but seven divides the number evenly. (Hint create four congruences and four lists of numbers).

Answer

49

- **5.3: Divisibility** by Harris Kwong is licensed CC BY-NC-SA 4.0.
- **5.7: Modular Arithmetic** by Harris Kwong is licensed CC BY-NC-SA 4.0.

CHAPTER OVERVIEW

6: Cryptography

Thumbnail: Alphabet shift ciphers are believed to have been used by Julius Caesar over 2,000 years ago. Caesar cipher with a shift of 3. Plaintext is at the top, Ciphertext is at the bottom. (Public Domain; Matt_Crypto via Wikipedia).

6.1: Cryptography

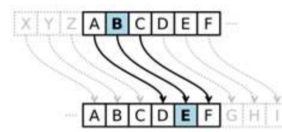
When people need to secretly store or communicate messages, they turn to cryptography. Cryptography involves using techniques to obscure a message so outsiders cannot read the message. It is typically split into two steps: encryption, in which the message is obscured, and decryption, in which the original message is recovered from the obscured form.

6.2: Substitution Ciphers

One simple encryption method is called a **substitution cipher**.

> **Substitution Cipher**
>
> A substitution cipher replaces each letter in the message with a different letter, following some established mapping.

A simple example of a substitution cipher is called the **Caesar cipher**, sometimes called a shift cipher. In this approach, each letter is replaced with a letter some fixed number of positions later in the alphabet. For example, if we use a shift of 3, then the letter A would be replaced with D, the letter 3 positions later in the alphabet. The entire mapping would look like: [1]



Original: ABCDEFGHIJKLMNOPQRSTUVWXYZ

Maps to: DEFGHIJKLMNOPQRSTUVWXYZABC

> **Example 1**
>
> Use the Caesar cipher with shift of 3 to encrypt the message: “We ride at noon”
>
> **Solution**
>
> We use the mapping above to replace each letter. W gets replaced with Z, and so forth, giving the encrypted message: ZH ULGH DW QRRQ.
>
> Notice that the length of the words could give an important clue to the cipher shift used. If we saw a single letter in the encrypted message, we would assume it must be an encrypted A or I, since those are the only single letters than form valid English words.
>
> To obscure the message, the letters are often rearranged into equal sized blocks. The message ZH ULGH DW QRRQ could be written in blocks of three characters as
>
> ZHU LGH DWQ RRQ.

> **Example 2**
>
> Decrypt the message GZD KNK YDX MFW JXA if it was encrypted using a shift cipher with shift of 5.
>
> **Solution**
>
> We start by writing out the character mapping by shifting the alphabet, with A mapping to F, five characters later in the alphabet.
>
> Original: ABCDEFGHIJKLMNOPQRSTUVWXYZ
>
> Maps to: FGHIJKLMNOPQRSTUVWXYZABCDE
>
> We now work backwards to decrypt the message. The first letter G is mapped to by B, so B is the first character of the original message. Continuing, our decrypted message is
>
> BUY FIF TYS HAR ESA.
>
> Removing spaces we get BUYFIFTYSHARESA. In this case, it appears an extra character was added to the end to make the groups of three come out even, and that the original message was “Buy fifty shares.”

> **Try it Now 1**
>
> Decrypt the message BNW MVX WNH if it was encrypted using a shift cipher with shift 9 (mapping A to J).
>
> **Answer**
>
> SEND MONEY

Notice that in both the ciphers above, the extra part of the alphabet wraps around to the beginning. Because of this, a handy version of the shift cipher is a cipher disc, such as the Alberti cipher disk shown here[2] from the 1400s. In a cipher disc, the inner wheel could be turned to change the cipher shift. This same approach is used for “secret decoder rings.”

The security of a cryptographic method is very important to the person relying on their message being kept secret. The security depends on two factors:

1. The security of the method being used
2. The security of the encryption key used

In the case of a shift cipher, the method is “a shift cipher is used.” The encryption key is the specific amount of shift used.

Suppose an army is using a shift cipher to send their messages, and one of their officers is captured by their enemy. It is likely the method and encryption key could become compromised. It is relatively hard to change encryption methods, but relatively easy to change encryption keys.

During World War II, the Germans’ Enigma encryption machines were captured, but having details on the encryption method only slightly helped the Allies, since the encryption keys were still unknown and hard to discover. Ultimately, the security of a message cannot rely on the method being kept secret; it needs to rely on the key being kept secret.

Encryption Security

The security of any encryption method should depend only on the encryption key being difficult to discover. It is not safe to rely on the encryption method (algorithm) being kept secret.

With that in mind, let’s analyze the security of the Caesar cipher.

Example 3

Suppose you intercept a message, and you know the sender is using a Caesar cipher, but do not know the shift being used. The message begins EQZP. How hard would it be to decrypt this message?

Solution

Since there are only 25 possible shifts, we would only have to try 25 different possibilities to see which one produces results that make sense. While that would be tedious, one person could easily do this by hand in a few minutes. A modern computer could try all possibilities in under a second.

Shift	Message	Shift	Message	Shift	Message	Shift	Message
1	DPYO	7	XJSI	13	RDMC	19	LXGW
2	COXN	8	WIRH	14	QCLB	20	KWFV
3	BNWM	9	VHQG	15	PBKA	21	JVEU
4	AMVL	10	UGPF	16	OAJZ	22	IUDT
5	ZLUK	11	TFOE	17	NZIY	23	HTCS
6	YKTJ	**12**	**SEND**	18	MYHX	24	GSBR
						25	FRAQ

In this case, a shift of 12 (A mapping to M) decrypts EQZP to SEND. Because of this ease of trying all possible encryption keys, the Caesar cipher is not a very secure encryption method.

Brute Force Attack

A brute force attack is a method for breaking encryption by trying all possible encryption keys.

To make a brute force attack harder, we could make a more complex substitution cipher by using something other than a shift of the alphabet. By choosing a random mapping, we could get a more secure cipher, with the tradeoff that the encryption key is harder to describe; the key would now be the entire mapping, rather than just the shift amount.

✔ Example 4

Use the substitution mapping below to encrypt the message “March 12 0300”

Original: $mathrm ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789$

Maps to: 2BQF5WRTD8IJ6HLCOSUVK3A0X9YZN1G4ME7P

Solution

Using the mapping, the message would encrypt to 62SQT ZN Y1YY

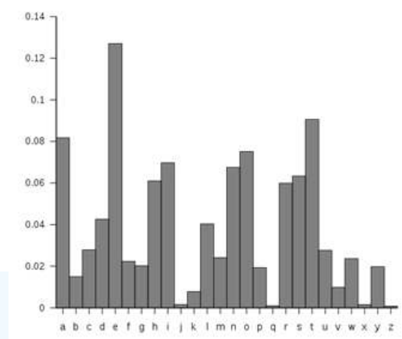
? Try it Now 2

Use the substitution mapping from Example 4 to decrypt the message C2SVX2VP

Answer

PARTY AT 9

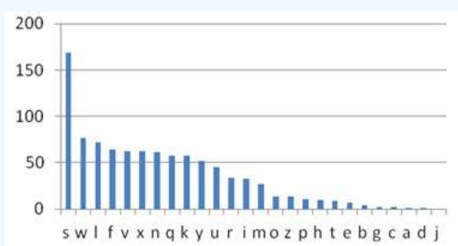
While there were only 25 possible shift cipher keys (35 if we had included numbers), there are about 10^{40} possible substitution ciphers[3]. That’s much more than a trillion trillions. It would be essentially impossible, even with supercomputers, to try every possible combination. Having a huge number of possible encryption keys is one important part of key security.



Unfortunately, this cipher is still not secure, because of a technique called frequency analysis, discovered by Arab mathematician Al-Kindi in the 9th century. English and other languages have certain letters than show up more often in writing than others.[4] For example, the letter E shows up the most frequently in English. The chart to the right shows the typical distribution of characters.

✔ Example 5

The chart to the right shows the frequency of different characters in some encrypted text. What can you deduce about the mapping?



Solution

Because of the high frequency of the letter S in the encrypted text, it is very likely that the substitution maps E to S. Since W is the second most frequent character, it likely that T or A maps to W. Because C, A, D, and J show up rarely in the encrypted text, it is likely they are mapped to from J, Q, X, and Z.

In addition to looking at individual letters, certain pairs of letters show up more frequently, such as the pair “th.” By analyzing how often different letters and letter pairs show up an encrypted message, the substitution mapping used can be deduced[5].

---

[1] en.Wikipedia.org/w/index.php?title=File:Caesar3.svg&page=1. PD

[2] en.Wikipedia.org/wiki/File:Alberti_cipher_disk.JPG

[3] There are 35 choices for what *A* maps to, then 34 choices for what *B* maps to, and so on, so the total number of possibilities is 35*34*33*…*2*1 = 35! = about 10^{40}

[4] en.Wikipedia.org/w/index.php?title=File:English_letter_frequency_(alphabetic).svg&page=1 PD

[5] For an example of how this is done, see en.Wikipedia.org/wiki/Frequency_analysis

6.3: Transposition Ciphers

Another approach to cryptography is **transposition cipher.**

Transposition Ciphers

A transposition cipher is one in which the order of characters is changed to obscure the message.

An early version of a transposition cipher was a Scytale[1], in which paper was wrapped around a stick and the message was written. Once unwrapped, the message would be unreadable until the message was wrapped around a same-sized stick again.

One modern transposition cipher is done by writing the message in rows, then forming the encrypted message from the text in the columns.

Example 6

Encrypt the message "Meet at First and Pine at midnight" using rows 8 characters long.

Solution

We write the message in rows of 8 characters each. Nonsense characters are added to the end to complete the last row.

MEETATFI

RSTANDPI

NEATMIDN

IGHTPXNR

We could then encode the message by recording down the columns. The first column, reading down, would be MRNI. All together, the encoded message would be MRNI ESEG ETAH TATT ANMP TDIX FPDN IINR. The spaces would be removed or repositioned to hide the size of table used, since that is the encryption key in this message.

Example 7

Decrypt the message CEE IAI MNL NOG LTR VMH NW using the method above with a table with rows of 5 characters.

Solution

Since there are total of 20 characters and each row should have 5 characters, then there will be $20/5 = 4$ rows.

We start writing, putting the first 4 letters, CEEI, down the first column.

CALLM

EINTH

EMORN

INGVW

We can now read the message: CALL ME IN THE MORNING VW. The VW is likely nonsense characters used to fill out the message.

More complex versions of this rows-and-column based transposition cipher can be created by specifying an order in which the columns should be recorded. For example, the method could specify that after writing the message out in rows that you should record the third column, then the fourth, then the first, then the fifth, then the second. This adds additional complexity that would make it harder to make a brute-force attack.

To make the encryption key easier to remember, a word could be used. For example, if the key word was "MONEY", it would specify that rows should have 5 characters each. The order of the letters in the alphabet would dictate which order to read the columns in. Since E, the 4th letter in the word, is the earliest letter in the alphabet from the word MONEY, the 4th column would be used first, followed by the 1st column (M), the 3rd column (N), the 2nd column (O), and the 5th column (Y).

Example 8

Encrypt the message BUY SOME MILK AND EGGS using a transposition cipher with key word MONEY.

Solution

Writing out the message in rows of 5 characters:

BUYSO
MEMIL
KANDE
GGSPK

We now record the columns in order 4 1 3 2 5:

SIDP BMKG YMNS UEAG OLEK

As before, we'd then remove or reposition the spaces to conceal evidence of the encryption key.

Try it Now 3

Encrypt the message "Fortify the embassy" using a transposition cipher with key word HELP

Answer

HELP gives column order 2 1 3 4.

FORT

IFYT

HEEM

BASS

YPLR

Encrypted text: OFE APF IHB YRY ESL TTM SR

To decrypt a keyword-based transposition cipher, we'd reverse the process. In the example above, the keyword MONEY tells us to begin with the 4th column, so we'd start by writing SIDP down the 4th column, then continue to the 1st column, 3rd column, etc.

Example 9

Decrypt the message RHA VTN USR EDE AIE RIK ATS OQR using a row-and-column transposition cipher with keyword PRIZED.

Solution

The keyword PRIZED tells us to use rows with 6 characters. Since D comes first in the alphabet, we start with 6th column. Since E is next in the alphabet, we'd follow with the 5th column. Continuing, the word PRIZED tells us the message was recorded with the columns in order 4 5 3 6 2 1.

For the decryption, we set up a table with 6 characters in each row. Since the beginning of the encrypted message came from the last column, we start writing the encrypted message down the last column.

					R
					H
					A
					V

The 5th column was the second one the encrypted message was read from, so is the next one we write to.

				T	R
				N	H
				U	A
				S	V

Continuing, we can fill out the rest of the message.

A	I	R	S	T	R
I	K	E	O	N	H
E	A	D	Q	U	A
R	T	E	R	S	V

Reading across the rows gives our decrypted message: AIRSTRIKEONHEADQUARTERSV

Unfortunately, since the transposition cipher does not change the frequency of individual letters, it is still susceptible to frequency analysis, though the transposition does eliminate information from letter pairs.

---

[1] en.Wikipedia.org/wiki/File:Skytala%26EmptyStrip-Shaded.png

6.4: Advanced shared symmetric-key methods

Both the substitution and transposition methods discussed so far are shared **symmetric-key** methods, meaning that both sender and receiver would have to have agreed upon the same secret encryption key before any methods could be sent.

All of the methods so far have been susceptible to frequency analysis since each letter is always mapped to the same encrypted character. More advanced methods get around this weakness. For example, the Enigma machines used in World War II had wheels that rotated. Each wheel was a substitution cipher, but the rotation would cause the substitution used to shift after each character.

For a simplified example, in the initial setup, the wheel might provide the mapping

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789}$
- Maps to: $\mathrm{2BQF5WRTD8IJ6HLCOSUVK3A0X9YZN1G4ME7P}$

After the first character is encrypted, the wheel rotates, shifting the mapping one space, resulting in a new shifted mapping:

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789}$
- Maps to: $\mathrm{P2BQF5WRTD8IJ6HLCOSUVK3A0X9YZN1G4ME7}$

Using this approach, no letter gets encrypted as the same character over and over.

Example 10

Encrypt the message “See me”. Use a basic Caesar cipher with shift 3 as the initial substitution, but shift the substitution one place after each character.

Solution

The initial mapping is

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$
- Maps to: $\mathrm{DEFGHIJKLMNOPQRSTUVWXYZABC}$

This would map the first letter, S to V. We would then shift the mapping by one.

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$
- Now maps to: $\mathrm{EFGHIJKLMNOPQRSTUVWXYZABCD}$

Now the next letter, E, will map to I. Again we shift the cipher

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$
- Now maps to: $\mathrm{FGHIJKLMNOPQRSTUVWXYZABCDE}$

The next letter, E, now maps to J. Continuing this process, the final message would be VIJSL.

Notice that frequency analysis is much less useful now, since the character E has been mapped to three different characters due to the shifting of the substitution mapping.

Try it Now 4

Decrypt the message KIQRV if it was encrypted using a basic Caesar cipher with shift 3 as the initial substitution, but shifting the substitution one place after each character.

Answer

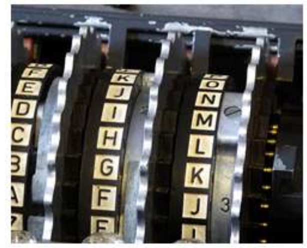
The initial mapping was:

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$
- Now maps to: $\mathrm{DEFGHIJKLMNOPQRSTUVWXYZABC}$

Using this, we can see the first character of the encrypted message, K, can be decrypted to the letter H. We now shift the mapping by one character.

- Original: $\mathrm{ABCDEFGHIJKLMNOPQRSTUVWXYZ}$
- Now maps to: $\mathrm{EFGHIJKLMNOPQRSTUVWXYZABCD}$

The second character in the message, I, can be decrypted to the letter E. Continuing this process of shifting and decrypting, KIQRV decrypts to HELLO.



The actual Engima machines used in WWII were more complex. Each wheel consisted of a complex substitution cipher, and multiple wheels were used in a chain[1]. The specific wheels used, order of the wheels, and starting position of the wheels formed the encryption key. While captured Engima devices provided the Allied forces details on the encryption method, the keys still had to be broken to decrypt messages.

These code breaking efforts led to the development of some of the first electronic computers by Alan Turing at Bletchley Park in the United Kingdom. This is generally considered the beginnings of modern computing[2].

In the 1970s, the U.S. government had a competition and ultimately approved an algorithm deemed DES (Data Encryption Standard) to be used for encrypting government data. It became the standard encryption algorithm used. This method used a combination of multiple substitution and transposition steps, along with other steps in which the encryption key is mixed with the message. This method uses an encryption key with length 56 bits, meaning there are 2^{56} possible keys.

This number of keys make a brute force attack extremely difficult and costly, but not impossible. In 1998, a team was able to find the decryption key for a message in 2 days, using about $250,000 worth of hardware. However, the price and time will go down as computer power increases.

From 1997 to 2001 the government held another competition, ultimately adopting a new method, deemed AES (Advanced Encryption Standard). This method uses encryption keys with 128, 192, or 256 bits, providing up to 2^{256} possible keys, making brute force attacks essentially impossible.

---

[1] http://en.wikipedia.org/wiki/File:En...abet_rings.jpg

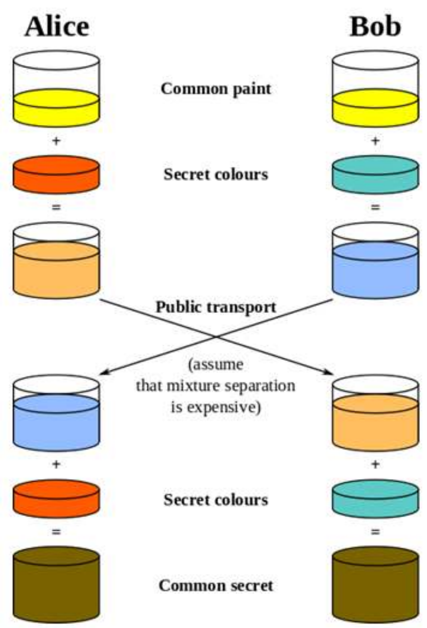
[2] For a good overview, see http://www.youtube.com/watch?v=5nK_ft0Lf1s

6.5: Public Key Cryptography

Suppose that you are connecting to your bank's website. It is possible that someone could intercept any communication between you and your bank, so you'll want to encrypt the communication. The problem is that all the encryption methods we've discussed require than both parties have already agreed on a shared secret encryption key. How can you and your bank agree on a key if you haven't already?

This becomes the goal of public key cryptography – to provide a way for two parties to agree on a key without a snooping third party being able to determine the key. The method relies on a one-way function; something that is easy to do one way, but hard to reverse. We will explore the Diffie-Hellman-Merkle key exchange method.

As an example, let's consider mixing paint. It's easy to mix paint to make a new color, but much harder to separate a mixed paint into the two original colors used.[1][2]



Using this analogy, Alice and Bob publically agree on a common starter color. Each then mixes in some of their own secret color. They then exchange their mixed colors.

Since separating colors is hard, even if a snooper were to obtain these mixed colors, it would be hard to obtain the original secret colors.

Once they have exchanged their mixed colors, Alice and Bob both add their secret color to the mix they obtained from the other person. In doing so, both Alice and Bob now have the same common secret color, since it contains a mix of the original common color, Alice's secret color, and Bob's secret color.

They now have a common secret color they can use as their encryption key, even though neither Alice nor Bob knows the other's secret color.

Likewise, there is no way for a snooper to obtain the common secret color without separating one of the mixed colors.

To get this process to work for computer communication, we need to have the process result in a share common number to act as the common secret encryption key. For this, we need a numerical one-way function.

Modular arithmetic

If you think back to doing division with whole numbers, you may remember finding the whole number result and the remainder after division.

Modulus [3]

The **modulus** is another name for the remainder after division.

For example, 17 mod 5 = 2, since if we divide 17 by 5, we get 3 with remainder 2.

Modular arithmetic is sometimes called clock arithmetic, since analog clocks wrap around times past 12, meaning they work on a modulus of 12. If the hour hand of a clock currently points to 8, then in 5 hours it will point to 1. While $8+5=13$, the clock wraps around after 12, so all times can be thought of as modulus 12. Mathematically, 13 mod 12 = 1.

Example 11

Compute: a) 10 mod 3 b) 15 mod 5 c) 2^7 mod 5

Solution

a) Since 10 divided by 3 is 3 with remainder 1,10 mod $3=1$

b) Since 15 divided by 5 is 3 with no remainder, 15 mod $5=0$

c) $2^7=128.128$ divide by 5 is 25 with remainder 3, so 2^7 mod $5=3$

Try it Now 5

Compute: a) 23 mod 7 b) 15 mod 7 c) 2034 mod 7

Answer

a) 2 b) 1 c) 4

Recall that when we divide 17 by 5, we could represent the result as 3 remainder 2, as the mixed number $3\frac{2}{5}$, or as the decimal 3.4. Notice that the modulus, 2, is the same as the numerator of the fractional part of the mixed number, and that the decimal part 0.4 is equivalent to the fraction . We can use these conversions to calculate the modulus of not-too-huge numbers on a standard calculator.

Modulus on a Standard Calculator

To calculate a mod n on a standard calculator

1. Divide a by n
2. Subtract the whole part of the resulting quantity
3. Multiply by n to obtain the modulus

Example 12

Calculate 31345 mod 419

Solution

$$\begin{array}{ll} 31345/419=74.8090692 & \text{Now subtract 74 to get just the decimal remainder} \\ 74.8090692-74=0.8090692 & \text{Multiply this by 419 to get the modulus} \\ 0.8090692\times 419=339 & \text{This tells us } 0.8090692 \text{ was equivalent to } \frac{339}{419} \end{array}$$

In the text above, only a portion of the decimal value was written down. In practice, you should try to avoid writing down the intermediary steps, and instead allow your calculator to retain as many decimal values as it can.

The one-way function

When you use a prime number p as a modulus, you can find a special number called a generator, g, so that g^n mod p will result in all the values from 1 to $p-1$

n	3^n	$3^n \mathrm{mod} 7$
1	3	3
2	9	2
3	27	6
4	81	4
5	243	5
6	729	1

In the table to the top, notice that when we give values of n from 1 to 6, we get out all values from 1 to 6. This means 3 is a generator when 7 is the modulus.

This gives us our one-way function. While it is easy to compute the value of g^n mod p when we know n, it is difficult to find the exponent n to obtain a specific value.

For example, suppose we use $p=23$ and $g=5$. If I pick n to be 6, I can fairly easily calculate $5^6 \bmod 23 = 15625 \bmod 23 = 8$

If someone else were to tell you $5^n \bmod 23 = 7$, it is much harder to find n. In this particular case, we'd have to try 22 different values for n until we found one that worked – there is no known easier way to find n other than brute-force guessing.

While trying 22 values would not take too long, when used in practice much larger values for p are used, typically with well over 500 digits. Trying all possibilities would be essentially impossible.

The key exchange

Before we can begin the key exchange process, we need a couple more important facts about modular arithmetic.

Modular Exponentialtion Rule

$\left(a^b \bmod n\right) = (a \bmod n)^b \bmod n$

Example 13

Compute 12^5 mod 7 using the exponentiation rule.

Solution

Evaluated directly: $12^5 = 248,832$, so 12^5 mod $7 = 248,823 \bmod 7 = 3$

Using the rule above, 12^5 mod $7 = (12 \bmod 7)^5 \bmod 7 = 5^5 \bmod 7 = 3125 \bmod 7 = 3$

You may remember a basic exponent rule from algebra:

$$\left(a^b\right)^c = a^{bc} = a^{cb} = (a^c)^b$$

For example:

$$64^2 = \left(4^3\right)^2 = 4^6 = \left(4^2\right)^3 = 16^3$$

We can combine the modular exponentiation rule with the algebra exponent rule to define the modular exponent power rule.

Modular Exponent Power Rule

$$\left(a^b \bmod n\right)^c \bmod n = \left(a^{bc} \bmod n\right) = \left(a^c \bmod n\right)^b \bmod n$$

Example 14

Verify the rule above if $a = 3, b = 4, c = 5,$ and $n = 7$

Solution

$\left(3^4 \bmod 7\right)^5 \bmod 7 = (81 \bmod 7)^5 \bmod 7 = 4^5 \bmod 7 = 1024 \bmod 7 = 2$

$\left(3^5 \bmod 7\right)^4 \bmod 7 = (243 \bmod 7)^4 \bmod 7 = 5^4 \bmod 7 = 625 \bmod 7 = 2,$ the same result.

Try it Now 6

Use the modular exponent rule to calculate $10000 \bmod 7$, by noting $10000 = 10^4$

Answer

$10000 \bmod 7 = 10^4 \bmod 7 = (10 \bmod 7)^4 \bmod 7 = 3^4 \bmod 7 = 81 \bmod 7 = 4$

This provides us the basis for our key exchange. While it will be easier to understand in the following example, here's the process:

1. Alice and Bob agree publically on values a prime p and generator g.
2. Alice picks some secret number a, while Bob picks some secret number b.
3. Alice computes $A = g^a \bmod p$ and sends it to Bob.
4. Bob computes $B = g^b \bmod p$ and sends it to Alice.
5. Alice computes $B^a \bmod p$, which is $\left(g^b \bmod p\right)^a \bmod p$.

6. Bob computes $A^b \bmod p$, which is $(g^a \bmod p)^b \bmod p$.

The modular exponent power rule tells us $(g^a \bmod p)^b \bmod p = (g^b \bmod p)^a \bmod p$, so Alice and Bob will arrive at the same shared value to use as a key, even though neither knows the other's secret number, and no eavesdropper can determine this value knowing only g, p, A, and B

Example 15

Alice and Bob publically share a generator and prime modulus. In this case, we'll use 3 as the generator and 17 as the prime.

Solution

	Alice		**Bob**
Alice and Bob publically share a generator and prime modulus.	$g = 3, p = 17$	Common info	$g = 3, p = 17$
Each then secretly picks a number n of their own.	$n = 8$	secret number	$n = 6$
Each calculates $g^n \bmod p$	$3^8 \bmod 17 = 16$		$3^6 \bmod 17 = 15$
They then exchange these resulting values.	$A = 16$ $B = 15$		$B = 15$ $A = 16$
Each then raises the value they received to the power of their secret $n \bmod p$	$B^n \bmod p =$ $15^8 \bmod 17 = 1$	mix in secret number	$A^n \bmod p =$ $16^6 \bmod 17 = 1$
The result is the shared secret key.	1	shared secret key	1

The shared secrets come out the same because of the modular exponent power rule $(a^b \bmod n)^c \bmod n = (a^c \bmod n)^b \bmod n$. Alice computed $(3^6 \bmod 17)^8 \bmod 17$ while Bob computed $(3^8 \bmod 17)^6 \bmod 17$, which the rule says will give the same results.

Notice that even if a snooper were to obtain both exchanged values $A = 16$ and $B = 15$, there is no way they could obtain the shared secret key from these without having at least one of Alice or Bob's secret numbers. There is no easy way to obtain the secret numbers from the shared values, since the function was a one-way function.

Using this approach, Alice and Bob can now use the shared secret key obtained as the key for a standard encryption algorithm like DES or AES.

Try it Now 7

Suppose you are doing a key exchange with Kylie using generator 5 and prime 23. Your secret number is 2. What number do you send to Kylie? If Kylie sends you the value 8, determine the shared secret key.

Answer

To compute the number we'd send to Kylie, we raise the generator to the power of our secret number modulus the prime: $5^2 \bmod 23 = 25 \bmod 23 = 2$.

If Kylie sends us the value 8, we determine the shared secret by raising her number to the power of our secret number modulus the prime: $8^2 \bmod 23 = 64 \bmod 23 = 18.18$would be the shared secret.

RSA

There are several other public-key methods used, including RSA, which is very commonly used. RSA involves distributing a public encryption key, which anyone can use to encrypt messages to you, but which can only be decrypted using a separate private key. You can think of this as sending an open padlock to someone – they can lock up information, but no one can unlock it without the key you kept secret.

RSA's security relies on the difficulty of factoring large numbers. For example, it's easy to calculate that 53 times 59 is 3127, but given the number 12,317 that is a product of two primes, it's much harder to find the numbers that multiply to give that number. It's exponentially harder when the primes each have 100 or more digits. Suppose we find two primes p and q and multiply them to get $n=pq$. This number will be very hard to factor. If we also know p and q, there are shortcuts to find two numbers e and d so that $m^{ed} \bmod n = m \bmod n$ for all numbers m. Without knowing the factorization of n, finding these values is very hard.

To use RSA, we generate two primes p and q and multiply them to get $n=pq$. since we know the factorization, we can easily find e and d so $m^{ed} = m \bmod n$. Now, we lock away p
q, and d. We then send the values e and n out publically. To encrypt a message m, the sender computes $S = m^e \bmod n$. As we saw earlier, the modulus is a one-way function which makes the original message very hard to recover from S. However, we have our private key d we can use to decrypt the message. When we receive the secret message S, we compute $S^d \bmod n = (m^e)^d \bmod n = m^{ed} \bmod n = m \bmod n$, recovering the original message[4].

Example 16

Suppose that Alice has computed $n = 3127, e = 3,$ and $d = 2011$. Show how Bob would encrypt the message 50 and how Alice would then decrypt it.

Solution

Bob would only know his message, $m = 50$ and Alice's public key: $n = 3127$ and $e = 3$. He would encrypt the message by computing $m^e \bmod n : 50^3 \bmod 3127 = 3047$

Alice can then decrypt this message using her private key d by computing $S^d \bmod n$:

$3047^{2011} \bmod 3127 = 50$

This method differs from Diffie-Hellman-Merkle because no exchange process is needed; Bob could send Alice an encrypted message using Bob's public key without having to communicate with Alice beforehand to determine a shared secret key. This is especially handy for applications like encrypting email, where both parties might not be online at the same time to perform a Diffie-Hellman-Merkle style key exchange.

---

[1] en.Wikipedia.org/w/index.php?...nge.svg&page=1

[2] For a video overview of this process, see http://www.youtube.com/watch?v=YEBfamv-_do

[3] Sometime, instead of seeing 17 mod 5 = 2, you'll see 17 ≡ 2 (mod 5). The ≡ symbol means "congruent to" and means that 17 and 2 are equivalent, after you consider the modulus 5.

[4] Many details have been left out, including how e and d are determined, and why this all works. For a bit more detail, see http://www.youtube.com/watch?v=wXB-V_Keiu8, or http://doctrina.org/How-RSA-Works-With-Examples.html

6.6: Other Secret Keeping Methods

While this chapter has focused on cryptography methods for keeping a message secure, there is another area of secret keeping called **steganography** which focuses on hiding the existence of a message altogether. Historically, steganography included techniques like invisible ink, watermarks, or embedding a secret message inside a longer document that appeared unimportant. The digital age has provided many new ways to hide messages in digital photos or in the background noise of a music file.

Resources

There are several calculation tools available based on the material in this chapter available at: http://www.opentextbookstore.com/mathinsociety/apps/

6.7: The One-Way Function

6.8: The key exchange

Substitution ciphers

In the questions below, if it specifies an alphabetic cipher, then the original map used letters only: ABCDEFGHIJKLMNOPQRSTUVWXYZ. If it specifies an alphanumeric cipher, then the original map used letters and numbers: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

1. Encrypt the message "SEND SUPPLIES" using an alphabetic Caesar cipher with shift 7 (mapping A to H).

2. Encrypt the message "CANCEL CONTRACT" using an alphanumeric Caesar cipher with shift 16 (mapping A to Q).

3. Decrypt the message "2R1 ONO 5SN OXM O" if it was encrypted using an alphanumeric Ceasar cipher with shift 10 (mapping A to K).

4. Decrypt the message "RJJY NSAJ SNHJ" if it was encrypted using an alphabetic Ceasar cipher with shift 5 (mapping A to F).

For questions 5-8 use this substitution mapping:

Original: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789
Maps to: HLCO2BQF5WRTZN1G4D8IJ6SUVK3A0X9YME7P

5. Use the substitution mapping to encrypt the message "DEAR DIARY"

6. Use the substitution mapping to encrypt the message "ATTACK AT SUNRISE"

7. Use the substitution mapping to decrypt the message "Z2DQ 2D1N"

8. Use the substitution mapping to decrypt the message "Z22 IHI3 YX3"

Transposition ciphers

9. Encrypt the message "Meet in the library at ten" using a tabular transposition cipher with rows of length 5 characters.

10. Encrypt the message "Fly surveillance over the northern county" using a tabular transposition cipher with rows of length 8 characters.

11. Decrypt the message "THE VHI NIE SAN SHT STI MQA DAN SDR S" if it was encrypted using a tabular transposition cipher with rows of length 7 characters.

12. Decrypt the message "DOLR UTIR INON KVEY AZ" if it was encrypted using a tabular transposition cipher with rows of length 6 characters.

13. Encrypt the message "Buy twenty million" using a tabular transposition cipher with the encryption keyword "RENT".

14. Encrypt the message "Attack from the northeast" using a tabular transposition cipher with the encryption keyword "POWER".

15. Decrypt the message "RYL OEN ONI TPM IEE YTE YDH WEA HRM S" if it was encrypted using a tabular transposition cipher with the encryption keyword "READING".

16. Decrypt the message "UYH SRT ABV HLN SEE L" if it was encrypted using a tabular transposition cipher with the encryption keyword "MAIL".

Shifting substitution ciphers

17. Encrypt the message "SEND SUPPLIES" using an alphabetic Caesar cipher that starts with shift 7 (mapping A to H), and shifts one additional space after each character is encoded.

18. Encrypt the message "CANCEL CONTRACT" using an alphabetic Caesar cipher that starts with shift 5 (mapping A to F), and shifts one additional space after each character is encoded.

Modular arithmetic

19. Compute

a. $15 \bmod 4$
b. $10 \bmod 5$

c. $257 \bmod 11$

20. Compute

a. $20 \bmod 4$
b. $14 \bmod 3$
c. $86 \bmod 13$

21. Determine if 4 is a generator modulus 11

22. Determine if 2 is a generator modulus 13

23. Use the modular exponent rule to calculate $157^{10} \bmod 5$

24. Use the modular exponent rule to calculate $133^{8} \bmod 6$

Diffie-Hellman-Merkle key exchange

25. Suppose you are doing a key exchange with Marc using generator 5 and prime 23. Your secret number is 7. Marc sends you the value 3. Determine the shared secret key.

26. Suppose you are doing a key exchange with Jen using generator 5 and prime 23. Your secret number is 4. Jen sends you the value 8. Determine the shared secret key.

RSA

27. Suppose that Alice has computed $n=33, e=7,$ and $d=3$. Show how Bob would encrypt the message 5 and how Alice would then decrypt it.

28. Suppose that Alice has computed $n=55, e=7,$ and $d=13$. Show how Bob would encrypt the message 8 and how Alice would then decrypt it.

Extensions

29. To further obscure a message, sometimes the usual alphabet characters are replaced with other symbols. Design a new set of symbols, and use it to encode a message. Exchange with a friend and see if they can decode your message.

30.To make an encryption harder to break, sometimes multiple substitution and transposition ciphers are used in sequence. For example, a method might specify that the first letter of the encryption keyword be used to determine the initial shift for a Caesar cipher (perhaps with a rotating cipher), and also be used for a transposition cipher. Design your own sequence of encryption steps and encrypt a message. Exchange with a friend and see if they can follow your process to decrypt the message.

31. When using large primes, computing values like $67^{24} \bmod 83$ can be difficult on a calculator without using additional tricks, since 67^{24} is a huge number. We will explore an approach used.

a. Notice that $67^{2} \bmod 83$ is fairly easy to calculate: $67^{2} \bmod 83=4489 \bmod 83=7$. Since $67^{4} \bmod 83=\left(67^{2}\right)^{2} \bmod 83$ can be rewritten using the modular exponent rule as $\left(67^{2} \bmod 83\right)^{2} \bmod 83$, this is also easy to evaluate: $67^{4} \bmod 83=\left(67^{2} \bmod 83\right)^{2} \bmod 83=7^{2} \bmod 83=49$ This process can be continued to find $67^{8} \bmod 83$ as $\left(6^{4}\right)^{2} \bmod 83$. Find this value, then find $67^{16} \bmod 83$ and $67^{32} \bmod 83$
b. There is a rule that $(a b) \bmod n=(a \bmod n)(b \bmod n) \bmod n$ Noting that $17000=170 \times 100$, calculate $17000 \bmod 83$ using the rule above.
c. Note that $67^{5}=67^{4} 67$. Use this, along with the rule from above and the results from part a to compute $67^{5} \bmod 83$
d. Note that $67^{7}=67^{4+2+1}=67^{4} 67^{2} 67^{1}$. Compute $67^{7} \bmod 83$.
e. Write 67^{24} as a product of powers of 67, and use this to compute $67^{24} \bmod 83$

32. Use the process from the previous question to evaluate $23^{34} \bmod 37$.

33. To encrypt text messages with RSA, the words are first converted into a string of numbers, and then encrypted. Several characters are usually combined together to produce a message number smaller than the modulus, but approximately the same size. Look up an ASCII table to convert the message “SCALE THE WALLS” to numbers, then encrypt it using the RSA public key $n=10823, e=5$. Since ASCII characters are two digits, pair up characters to form four-digit numbers before encoding. For example A is 65 and B is 66, so the character pair AB could be treated as the number 6566 and encrypted as 10148

34. Explore approaches to steganography that don't require specialized software. Attempt to hide a message using one of these techniques, and see if a fellow student can detect the message.

35. When you visit a secure website, your web browser will report that the site's identity has been verified by a third party, called a certificate authority. This is meant to assure you that you are visiting the actual company's website. Research how these certificates work.

CHAPTER OVERVIEW

7: Algebraic Models

7.1: Linear Equations

Populations of people, animals, and items are growing all around us. By understanding how things grow, we can better understand what to expect in the future. In this chapter, we focus on time-dependent change.

Graphing a Line from an Equation

Equations whose graphs are straight lines are called **linear equations.** The following are some examples of linear equations:

$$2x-3y=6, \quad 3x=4y-7, \quad y=2x-5, \quad 2y=3, \quad \text{and} \quad x-2=0$$

A line is completely determined by two points. Therefore, to graph a linear equation we need to find the coordinates of two points. This can be accomplished by choosing an arbitrary value for x or y and then solving for the other variable.

Example 7.1.1

Graph the line: $y=3x+2$

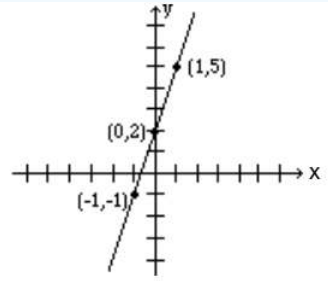
Solution

We need to find the coordinates of at least two points. We arbitrarily choose x = - 1, x = 0, and x = 1.

- If x = -1, then y = 3(-1) + 2 or -1. Therefore, (-1, -1) is a point on this line.
- If x = 0, then y = 3(0) + 2 or y = 2. Hence the point (0, 2).
- If x = 1, then y = 5, and we get the point (1, 5).

Below, the results are summarized, and the line is graphed.

x	-1	0	1
y	-1	2	5



Example 7.1.2

Graph the line: $2x+y=4$

Solution

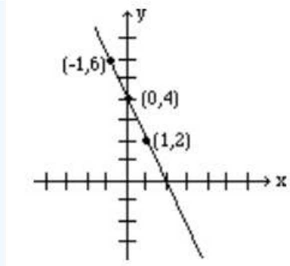
Again, we need to find coordinates of at least two points.

We arbitrarily choose x = -1, x = 0, and y = 2.

- If x = -1, then 2(-1) + y = 4 which results in y = 6. Therefore, (-1, 6) is a point on this line.
- If x = 0, then 2(0) + y = 4, which results in y = 4. Hence the point (0, 4).
- If y = 2, then 2x + 2 = 4, which yields x = 1, and gives the point (1, 2).

The table below shows the points, and the line is graphed.

x	-1	0	1
y	6	4	2



Intercepts

The points at which a line crosses the coordinate axes are called the **intercepts**.
When graphing a line by plotting two points, using the intercepts is often preferred because they are easy to find.

- To find the value of the x-intercept, we let y = 0
- To find the value of the y-intercept, we let x = 0.

Example 7.1.3

Find the intercepts of the line: $2x-3y=6$, and graph.

Solution

To find the x-intercept, let y = 0 in the equation, and solve for x.

$$\begin{aligned} 2x-3(0) &= 6 \\ 2x-0 &= 6 \\ 2x &= 6 \\ x &= 3 \end{aligned}$$

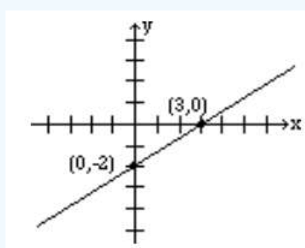
Therefore, the x-intercept is the point (3,0).

To find the y-intercept, let x = 0 in the equation, and solve for y.

$$\begin{aligned} 2(0)-3y &= 6 \\ 0-3y &= 6 \\ -3y &= 6 \\ y &= -2 \end{aligned}$$

Therefore, the y-intercept is the point (0, -2).

To graph the line, plot the points for the x-intercept (3,0) and the y-intercept (0, -2), and use them to draw the line.



Horizontal and Vertical Lines

When an equation of a line has only one variable, the resulting graph is a horizontal or a vertical line.

- The graph of the line $x=a$, where a is a constant, is a vertical line that passes through the point $(a,0)$. Every point on this line has the x-coordinate equal to a, regardless of the y-coordinate.

- The graph of the line $y = b$, where b is a constant, is a horizontal line that passes through the point $(0, b)$. Every point on this line has the y-coordinate equal to b, regardless of the x-coordinate.

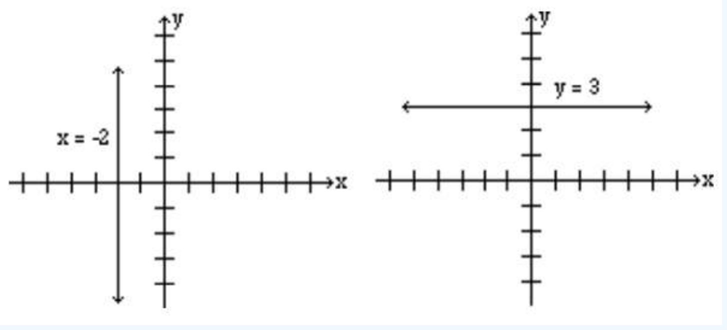
Example 7.1.4

Graph the lines: x = -2 , and y = 3.

Solution

The graph of the line x = -2 is a vertical line that has the x-coordinate -2 no matter what the y-coordinate is. The graph is a vertical line passing through point (-2, 0).

The graph of the line y = 3, is a horizontal line that has the y-coordinate 3 regardless of what the x-coordinate is. Therefore, the graph is a horizontal line that passes through point (0, 3).



Note: Most students feel that the coordinates of points must always be integers. This is not true, and in real life situations, not always possible. Do not be intimidated if your points include numbers that are fractions or decimals.

Slope

So far we have learned how to graph a line by choosing two points on the line. The graph of a line can also be determined if one point and the "steepness" of the line is known. The number that refers to the steepness or inclination of a line is called the **slope** of the line.

From previous math courses, you may remember slope as the "rise over run," or "the vertical change over the horizontal change" and have often seen it expressed as:

$$\frac{\text{rise}}{\text{run}}, \frac{\text{vertical change}}{\text{horizontal change}}, \frac{\Delta y}{\Delta x} \text{ etc.}$$

Definition: Slope

If (x_1, y_1) and (x_2, y_2) are two different points on a line, the **slope** of the line is

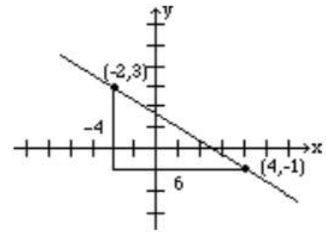
$$\text{slope} = m = \frac{y_2 - y_1}{x_2 - x_1} \tag{7.1.1}$$

Example 7.1.5

Find the slope of the line passing through points (−2, 3) and (4, −1), and graph the line.

Solution

Let (x_1, y_1) = (−2, 3) and (x_2, y_2) = (4, −1), then the slope (via Equation 7.1.1) is

$m = \frac{-1-3}{4-(-2)} = \frac{-4}{6} = -\frac{2}{3}$	

To give the reader a better understanding, both the vertical change, -4, and the horizontal change, 6, are shown in the above figure.

When two points are given, it does not matter which point is denoted as (x_1, y_1) and which (x_2, y_2). The value for the slope will be the same.

In Example $7.1.5$, if we instead choose (x_1, y_1) = (4, −1) and (x_2, y_2) = (−2, 3), then we will get the same value for the slope as we obtained earlier.

The steps involved are as follows.

$$m = \frac{3-(-1)}{-2-4} = \frac{4}{-6} = -\frac{2}{3}$$

The student should further observe that

- if a line rises when going from left to right, then it has a positive slope. In this situation, as the value of x increases, the value of y also increases
- if a line falls going from left to right, it has a negative slope; as the value of x increases, the value of y decreases.

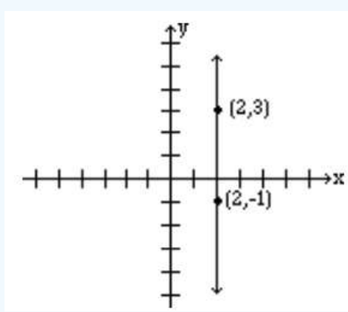
✓ Example 7.1.6

Find the slope of the line that passes through the points (2, 3) and (2, −1), and graph.

Solution

Let (x_1, y_1) = (2, 3) and (x_2, y_2) = (2, −1) then the slope is

$$m = \frac{-1-3}{2-2} = \frac{4}{0} = \text{undefined.}$$



Note: The slope of a vertical line is undefined.

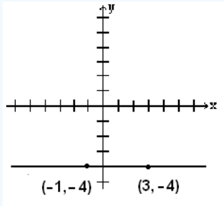
✓ Example 7.1.7

Find the slope of the line that passes through the points (−1, −4) and (3, −4)

Solution

Let (x_1, y_1) = (−1, −4) and (x_2, y_2) = (3, −4), then the slope is

$$m=\frac{-4-(-4)}{3-(-1)}=\frac{0}{4}=0$$



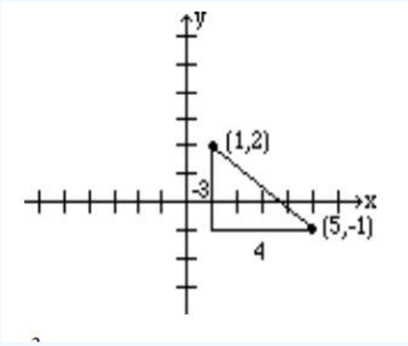
Note: The slope of a horizontal line is 0

Example 7.1.8

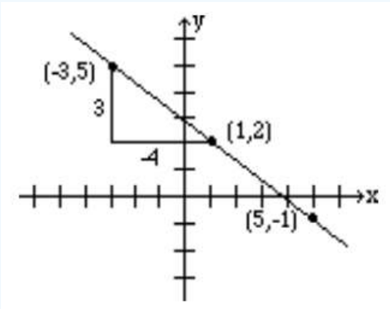
Graph the line that passes through the point (1, 2) and has slope $-\frac{3}{4}$.

Solution

Slope equals $\frac{\text{rise}}{\text{run}}$. The fact that the slope is $\frac{-3}{4}$, means that for every rise of -3 units (fall of 3 units) there is a run of 4. So if from the given point (1, 2) we go down 3 units and go right 4 units, we reach the point (5, -1). The graph is obtained by connecting these two points.



Alternatively, since $\frac{3}{-4}$ represents the same number, the line can be drawn by starting at the point (1,2) and choosing a rise of 3 units followed by a run of -4 units. So from the point (1, 2), we go up 3 units, and to the left 4, thus reaching the point (-3, 5) which is also on the same line. See figure below.



Example 7.1.9

Find the slope of the line $2x+3y=6$.

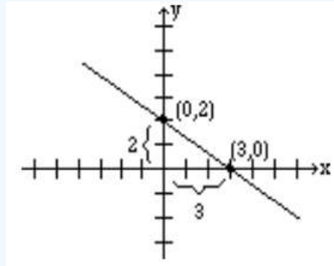
Solution

In order to find the slope of this line, we will choose any two points on this line.

Again, the selection of x and y intercepts seems to be a good choice. The x-intercept is (3, 0), and the y-intercept is (0, 2). Therefore, the slope is

$$m=\frac{2-0}{0-3}=-\frac{2}{3}. \tag{7.1.2}$$

The graph below shows the line and the x-intercepts and y-intercepts:



Example 7.1.10

Find the slope of the line $y=3x+2$.

Solution

We again find two points on the line, e.g., (0, 2) and (1, 5). Therefore, the slope is

$$m=\frac{5-2}{1-0}=\frac{3}{1}=3.$$

Look at the slopes and the y-intercepts of the following lines.

The line	slope	y-intercept
$y=3x+2$	3	2
$y=-2x+5$	-2	5
$y=\frac{3}{2}x-4$	$\frac{3}{2}$	-4

It is no coincidence that when an equation of the line is solved for y, the coefficient of the x term represents the slope, and the constant term represents the y-intercept.

In other words, for the line $y=mx+b$, m is the slope, and b is the y-intercept.

Example 7.1.11

Determine the slope and y-intercept of the line $2x+3y=6$.

Solution

We solve for y:

$$2x+3y=6$$
$$3y=-2x+6$$
$$y=(-2/3)x+2$$

The slope = the coefficient of the x term = − 2/3.

The y-intercept = the constant term = 2.

So far, we were given an equation of a line and were asked to give information about it. For example, we were asked to find points on the line, find its slope and even find intercepts. Now we are going to reverse the process. That is, we will be given either two points, or a point and the slope of a line, and we will be asked to find its equation.

An equation of a line can be written in three forms, the **slope-intercept form**, the **point-slope form,** or the **standard form**. We will discuss each of them in this section.

A line is completely determined by two points, or by a point and slope. The information we are given about a particular line will influence which form of the equation is most convenient to use. Once we know any form of the equation of a line, it is easy to re-express the equation in the other forms if needed.

The Slope-Intercept Form of a Line: $y = mx + b$

In the last section we learned that the equation of a line whose slope = m and y-intercept = b is

$$\mathbf{y} = \mathbf{mx} + \mathbf{b}. \tag{7.1.3}$$

This is called the **slope-intercept form** of the line and is the most commonly used form.

Example 7.1.12

Find an equation of a line whose slope is 5, and y-intercept is 3.

Solution

Since the slope is $m = 5$, and the y- intercept is $b = 3$, the equation is $y = 5x + 3$.

Example 7.1.13

Find the equation of the line that passes through the point $(2, 7)$ and has slope 3.

Solution

Since $m = 3$, the partial equation is $y = 3x + b$.

Now b can be determined by substituting the point $(2, 7)$ in the equation $y = 3x + b$.

$$\begin{aligned} 7 &= 3(2) + b \\ b &= 1 \end{aligned}$$

Therefore, the equation is $y = 3x + 1$.

Example 7.1.14

Find an equation of the line that passes through the points (-1, 2), and (1, 8).

Solution

$m = \frac{8-2}{1-(-1)} = \frac{6}{2} = 3$. So the partial equation is $y = 3x + b$.

We can use either of the two points (-1, 2) or (1, 8), to find b. Substituting (-1, 2) gives

$$\begin{aligned} 2 &= 3(-1) + b \\ 5 &= b \end{aligned}$$

So the equation is $y = 3x + 5$.

Example 7.1.15

Find an equation of the line that has x-intercept 3, and y-intercept 4.

Solution

x-intercept = 3, and y-intercept = 4 correspond to the points (3, 0), and (0, 4), respectively.

$$m = \frac{4-0}{0-3} = -\frac{4}{3}$$

We are told the y-intercept is 4; thus b = 4

Therefore, the equation is $y = -\frac{4}{3}x + 4$.

The Point-Slope Form of a Line: $y - y_1 = m(x - x_1)$

The **point-slope** form is useful when we know two points on the line and want to find the equation of the line.

Let L be a line with slope m, and known to contain a specific point (x_1, y_1). If (x, y) is any other point on the line L, then the definition of a slope leads us to the **point-slope form** or point-slope formula.

The slope is $\frac{y-y_1}{x-x_1} = m$

Multiplying both sides by $(x - x_1)$ gives the point-slope form:

$$\mathbf{y - y_1 = m(x - x_1)} \tag{7.1.4}$$

Example 7.1.16

Find the point-slope form of the equation of a line that has slope 1.5 and passes through the point (12,4).

Solution

Substituting the point $(x_1, y_1) = (12, 4)$ and $m = 1.5$ in the point-slope formula, we get

$$\mathbf{y - y_1 = m(x - x_1)}$$

$$y - 4 = 1.5(x - 12)$$

The student may be tempted to simplify this into the slope intercept form $y = mx + b$. But since the problem specifically requests **point-slope** form we will not simplify it.

The Standard Form of a Line: $Ax + By = C$

Another useful form of the equation of a line is the standard form.

If we know the equation of a line in point-slope form, $y - y_1 = m(x - x_1)$, or if we know the equation of the line in slope-intercept form $y = mx + b$, we can simplify the formula to have all terms for the x and y variables on one side of the equation, and the constant on the other side of the equation.

The result is referred to as the **standard form** of the line:

$$\mathbf{Ax + By = C}. \tag{7.1.5}$$

Example 7.1.17

Using the point-slope formula, find the standard form of an equation of the line that passes through the point (2, 3) and has slope $-3/5$.

Solution: Substituting the point (2, 3) and $m = -3/5$ in the point-slope formula, we get

$$y - 3 = -3/5(x - 2)$$

Multiplying both sides by 5 gives us

$$\begin{aligned} &5(y-3) = -3(x-2) \\ &5y - 15 = -3x + 6 \\ &3x + 5y = 21 \quad \text{Standard Form} \end{aligned}$$

Example 7.1.18

Find the standard form of the line that passes through the points (1, -2), and (4, 0).

Solution

First we find the slope: $m = \frac{0-(-2)}{4-1} = \frac{2}{3}$

Then, the point-slope form is: $y - (-2) = \frac{2}{3}(x - 1)$

Multiplying both sides by 3 gives us

$$\begin{aligned}&3(y+2)=2(x-1)\\&3y+6=2x-2\\&-2x+3y=-8\\&2x-3y=8\quad \text{Standard Form}\end{aligned}$$

We should always be able to convert from one form of an equation to another. For example, if we are given a line in the slope-intercept form, we should be able to express it in the standard form, and vice versa.

✔ Example 7.1.19

Write the equation $y=-\frac{2}{3}x+3$ in the standard form.

Solution

Multiplying both sides of the equation by 3, we get

$$\begin{aligned}&3y=-2x+9\\&2x+3y=9\quad \text{Standard Form}\end{aligned}$$

✔ Example 7.1.20

Write the equation $3x-4y=10$ in the slope-intercept form.

Solution

Solving for y, we get

$$\begin{aligned}&-4y=-3x+10\\&y=\frac{3}{4}x-\frac{5}{2}\quad \text{Standard Form}\end{aligned}$$

Finally, we learn a very quick and easy way to write an equation of a line in the standard form. But first we must learn to find the slope of a line in the standard form by inspection.

By solving for y, it can easily be shown that the slope of the line $Ax+By=C$ is $-A/B$.
The reader should verify this.

✔ Example 7.1.21

Find the slope of the following lines, by inspection.

a. $3x-5y=10$
b. $2x+7y=20$
c. $4x-3y=8$

Solution

a. $A=3$, $B=-5$, therefore, $m=\frac{-3}{-5}=\frac{3}{5}$
b. $A=2$, $B=7$, therefore, $m=\frac{-2}{7}=-\frac{2}{7}$
c. $m=\frac{-4}{-3}=\frac{4}{3}$

Now that we know how to find the slope of a line in the standard form by inspection, our job in finding the equation of a line is going to be easy.

Example 7.1.22

Find an equation of the line that passes through (2, 3) and has slope - 4/5.

Solution

Since the slope of the line is - 4/5, we know that the left side of the equation is $4x+5y$, and the partial equation is going to be

$$4x+5y=c$$

Of course, c can easily be found by substituting for x and y.

$$\begin{array}{l}4(2)+5(3)=c \\ 23=c\end{array}$$

The desired equation is

$$4x+5y=23.$$

If you use this method often enough, you can do these problems very quickly.

An Application Problem

Now that we have learned how to write the equation of a line, we get to apply these ideas in a variety of real-life situations.

Read the problem carefully. Highlight important information. Keep track of which values correspond to the independent variable (x) and which correspond to the dependent variable (y).

Example 7.1.1

A taxi service charges $0.50 per mile plus a $5 flat fee. What will be the cost of traveling 20 miles? What will be cost of traveling x miles?

Solution

x = distance traveled, in miles and y = cost in dollars

The cost of traveling 20 miles is

$$y=(0.50)(20)+5=10+5=15$$

The cost of traveling x miles is

$$y=(0.50)(x)+5=0.50x+5$$

In this problem, $0.50 per mile is referred to as the **variable cost**, and the flat charge $5 as the **fixed cost**. Now if we look at our cost equation $y=.50x+5$, we can see that the variable cost corresponds to the slope and the fixed cost to the y-intercept.

We summarize the forms for the equation of a line below:

Slope Intercept form: $\mathbf{y=mx+b}$,
where m = slope, b = y-intercept

Point Slope form: $\mathbf{y-y_1=m(x-x_1)}$,
where m = slope, (x_1, y_1) is a point on the line

Standard form: $\mathbf{Ax+By=C}$

Horizontal Line: $\mathbf{y=b}$
where b = y-intercept

Vertical Line: $\mathbf{x=a}$
where a = x-intercept

- **8.1: Introduction** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.
- **1.1: Graphing a Linear Equation** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.
- **1.2: Slope of a Line** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.
- **1.3: Determining the Equation of a Line** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.
- **1.4: Applications** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.

7.2: Modeling with Linear Equations

In this section, you will learn to use linear functions to model real-world applications

Read the problem carefully. Highlight important information. Identify what each variable represents in the context of the problem.

Example 7.2.1

It costs $750 to manufacture 25 items, and $1000 to manufacture 50 items. Assuming a linear relationship holds, find the cost equation, and use this equation to predict the cost of 100 items.

Solution

We let x = the number of items manufactured, and let y = the cost.

Solving this problem is equivalent to finding an equation of a line that passes through the points (25, 750) and (50, 1000).

$$m=\frac{1000-750}{50-25}=10$$

Therefore, the partial equation is $y=10x+b$

By substituting one of the points in the equation, we get $b=500$

Therefore, the cost equation is $y=10x+500$. This equation is the linear model for this problem.

Now use the linear model to find the cost of 100 items. Substitute $x=100$ in the equation $y=10x+500$

So the cost is

$$y=10(100)+500=1500$$

It costs $1500 to manufacture 100 items.

Example 7.2.2

The freezing temperature of water in Celsius is 0 degrees and in Fahrenheit 32 degrees. And the boiling temperatures of water in Celsius, and Fahrenheit are 100 degrees, and 212 degrees, respectively. Write a conversion equation from Celsius to Fahrenheit and use this equation to convert 30 degrees Celsius into Fahrenheit.

Solution

Let us look at what is given.

Celsius	Fahrenheit
0	32
100	212

We let C = the degrees in Celsius, and let F = the degrees in Fahrenheit.

Again, solving this problem is equivalent to finding an equation of a line that passes through the points (0, 32) and (100, 212).

Since we are finding a linear relationship, we are looking for an equation $y=mx+b$, or in this case $F=mC+b$, where x or C represent the temperature in Celsius, and y or F the temperature in Fahrenheit.

$$\text{slope m }=\frac{312-32}{100-0}=\frac{9}{5}$$

The equation is $F=\frac{9}{5}C+b$

Substituting the point (0, 32), we get b = 32 and the conversion equation is

$$F=\frac{9}{5}C+32.$$

To convert 30 degrees Celsius into Fahrenheit, substitute $C = 30$ in the equation

$$\mathrm{F} = \frac{9}{5}\mathrm{C} + 32$$
$$\mathrm{F} = \frac{9}{5}(30) + 32 = 86$$

Thus, 30 degrees Celsius is equal to 86 degrees Fahrenheit.

Example 7.2.3

The variable cost to manufacture a product is $10 per item and the fixed cost $2500. If x represents the number of items manufactured and y represents the total cost, write the cost function.

Solution

- The variable cost of $10 per item tells us that $m = 10$.
- The fixed cost represents the y-intercept. So $b = 2500$.

Therefore, the cost function is $y = 10x + 2500$.

Example 7.2.4

Assume a car depreciates by the same amount each year. Joe purchased a car in 2010 for $16,800. In 2014 it is worth $12,000. Find the linear model. Use this model to predict how much the car will be worth in 2020.

Solution:

We let x = the number of years after 2010, and let y = the cost.

Solving this problem is equivalent to finding an equation of a line that passes through the points (0, 16800) and (4, 12000).

To find the linear model for this problem, we need to find the slope.

$$m = \frac{12000 - 16800}{4 - 0} = -1200$$

The slope indicates that the rate of depreciation each year is $-1200. Thus, the linear model for this problem is: $y = -1200x + 16,800$
Now, to find out how much the car will be worth in 2020, we need to know how many years that is from the purchase year. Since it is ten years later, $x = 10$.

$$y = -1200(10) + 16,800 = -12,000 + 16,800 = 4,800 \tag{7.2.1}$$

The car will be worth $4800 in 2020.

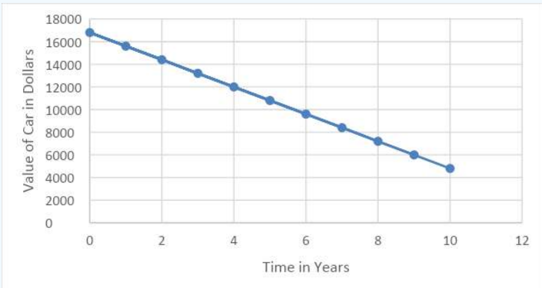


Figure 7.2.4: Graph of Car Value Depreciation

Note: The value of the car over time follows a decreasing straight line.

Example 7.2.5

The cost y, in dollars, of a gym membership for n months can be described by the linear model $y = 30n + 70$. What does this model tell us?

Solution

The value for y when $n = 0$ in this equation is 70, so the initial starting cost is $70. This tells us that there must be an initiation or start-up fee of $70 to join the gym.

The value for the slope, m in the equation is 30, so the cost increases by $30 each month. This tells us that the monthly membership fee for the gym is $30 a month.

Example 7.2.6

The population of Canada in the year 1980 was 24.5 million, and in the year 2010 it was 34 million. The population of Canada over that time period can be approximately modeled by a linear function. Let x represent time as the number of years after 1980 and let y represent the size of the population.

a. Write the linear function that gives a relationship between the time and the population.
b. Assuming the population continues to grow linearly in the future, use this equation to predict the population of Canada in the year 2025.

Solution

The problem can be made easier by using 1980 as the base year, that is, we choose the year 1980 as the year zero. This will mean that the year 2010 will correspond to year 30. Now we look at the information we have:

Year	Population
0 (1980)	24.5 million
30 (2010)	34 million

a. Solving this problem is equivalent to finding an equation of a line that passes through the points (0, 24.5) and (30, 34). We use these two points to find the slope:

$$m = \frac{34 - 24.5}{30 - 0} = \frac{9.5}{30} = 0.32$$

The y-intercept occurs when $x = 0$, so $b = 24.5$. We write the linear model

$$y = 0.32x + 24.5$$

b. Now to predict the population in the year 2025, we let $x = 2025 - 1980 = 45$

$$y = 0.32x + 24.5$$
$$y = 0.32(45) + 24.5 = 38.9$$

In the year 2025, we predict that the population of Canada will be 38.9 million people.

Note that we assumed the population trend will continue to be linear. Therefore if population trends change and this assumption does not continue to be true in the future, this prediction may not be accurate.

Definition: Linear Growth

A quantity grows linearly if it grows by a constant amount for each unit of time.

Example 7.2.7: City Growth

Suppose in Flagstaff Arizona, the number of residents increased by 1000 people per year. If the initial population was 46,080 in 1990, can you predict the population in 2013? This is an example of linear growth because the population grows by a constant amount. We list the population in future years below by adding 1000 people for each passing year.

	1990	1991	1992	1993	1994	1995	1996
Year	0	1	2	3	4	5	6
Population	46,080	47,080	48,080	49,080	50,080	51,080	52,080

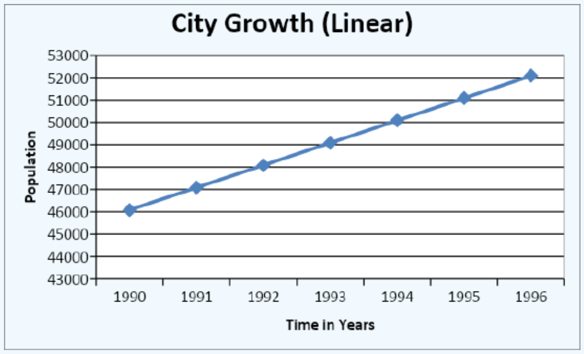


Figure 7.2.7: Graph of Linear Population Growth

Solution:

The population growth, *y*, can be modeled with a linear equation. The initial population is 46,080. The future population depends on the number of years, *t*, after the initial year. The model is $y = 1000t + 46,080$. Note, we chose to use the variable *t* as a simple reminder that t represents *time*. We could continue to use the variable *x*, or any other letter for that matter, but *t for time* makes sense.

To predict the population in 2013, we identify how many years it has been from 1990 (which is year zero). So *t* = 23 for the year 2013.

$$y = 1000(23) + 46,080 = 69,080 \tag{7.2.2}$$

The population of Flagstaff in 2013 will be 69,080 people.

Example 7.2.8: Antique Frog Collection

Dora has inherited a collection of 30 antique frogs. Each year she vows to buy two frogs a month to grow the collection. This is an additional 24 frogs per year. How many frogs will she have in six years? How long will it take her to reach 510 frogs?

Solution

The initial population is 30 frogs, so b=30. The rate of change is 24 frogs per year, so m=24. The linear growth model for this problem is:

$$y = 24t + 30 \tag{7.2.3}$$

where t = time in years and y = the number of frogs

The first question asks how many frogs will Dora have in six years so, *t* = 6.

$$y = 24(6) + 30 = 144 + 30 = 174 \tag{7.2.4}$$

frogs.

The second question asks for the time it will take for Dora to collect 510 frogs. So, $y = 510$and we will solve for *t*.

$$\begin{aligned}510 &= 24t+30\\480 &= 24t\\20 &= t\end{aligned}$$

It will take 20 years to collect 510 antique frogs.

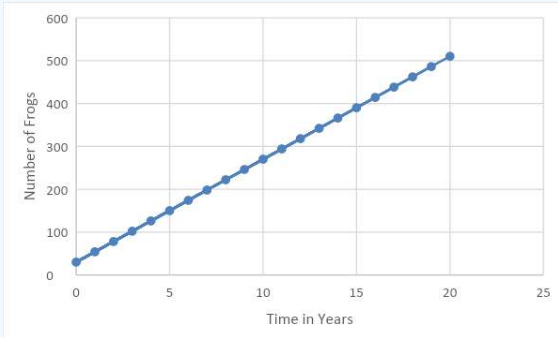


Figure 7.2.8: **Graph of Antique Frog Collection**

Note: The graph of the number of antique frogs Dora accumulates over time follows a straight line.

Try it Now 1

The number of stay-at-home fathers in Canada has been growing steadily[1]. While the trend is not perfectly linear, it is fairly linear. Use the data from 1976 and 2010 to find an explicit formula for the number of stay-at-home fathers, then use it to predict the number of stay-at-home fathers in 2020.

Year	1976	1984	1991	2000	2010
Number of stay-at-home fathers	20,610	28,725	43,530	47,665	53,555

Answer

We let t = the number of years after 1976, and let y = the number of stay-at-home fathers.

From the table we know that 1976 corresponds to $t=0$ and the number of stay-at-home fathers is $y=20,610$.

From 1976 to 2010 the number of stay-at-home fathers increased by

$53,555-20,610=32,945$

This happened over 34 years, so the rate of change (slope) is $32,945/34=969$

$y=969t+20,610$

Predicting for 2020, we use $t=44$

$y=969(44)+20,610=63,246$

There will be 63,246 stay-at-home fathers in 2020.

[1] www.fira.ca/article.php?id=140

7.3: Modeling with Quadratic Equations

Content not used for this course.

7.4: Exponential Growth

The next growth we will examine is exponential growth. Linear growth occurs by adding the same amount in each unit of time. Exponential growth happens when an initial population increases by the same percentage or factor over equal time increments or generations. This is known as relative growth and is usually expressed as percentage.

Suppose that every year, only 10% of the fish in a lake have surviving offspring. If there were 100 fish in the lake last year, there would now be 110 fish. If there were 1000 fish in the lake last year, there would now be 1100 fish. Absent any inhibiting factors, populations of people and animals tend to grow by a percent of the existing population each year.

Suppose our lake began with 1000 fish, and 10% of the fish have surviving offspring each year. We start with the initial population, 1000 fish, $P = 1000$. How do we calculate the population after 1 year? The new population, A, will be the initial population plus an additional 10%. Symbolically:

$A = P + 0.10P$

Notice this could be condensed to a shorter form by factoring:

$A = P + 0.10P = P(1 + 0.10) = P(1.10)$

While 10% is the **growth rate**, 1.10 is the **growth multiplier**. Notice that 1.10 can be thought of as "the original 100% plus an additional 10%"

To calculate the fish population after 1 year,

$A = P(1.10) = 1000(1.10) = 1100$

We could then calculate the population in the next years by using the population from the previous year:

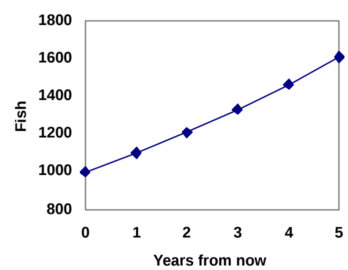
$A = 1100(1.10) = 1210$

$A = 1210(1.10) = 1331$

Notice that in the first year, the population grew by 100 fish, in the second year, the population grew by 110 fish, and in the third year the population grew by 121 fish.

While there is a constant percentage growth, the actual increase in number of fish is increasing each year.

Graphing these values we see that this growth doesn't quite appear linear.



To get a better picture of how this percentage-based growth affects things, we need an explicit form, so we can quickly calculate values further out in the future.

Like we did for the linear model, we will start building from the recursive equation:

For year 1: $A = 1000(1.10)$

For year 2: $A = 1000(1.10)(1.10)) = 1000(1.10)^2$

For year 3: $A = 1000(1.10)(1.10)(1.10) = 1000(1.10)^3$

For year 4: $A = 1000(1.10)(1.10)(1.10)(1.10) = 1000(1.10)^4$

Observing a pattern, we can generalize the explicit form to be:

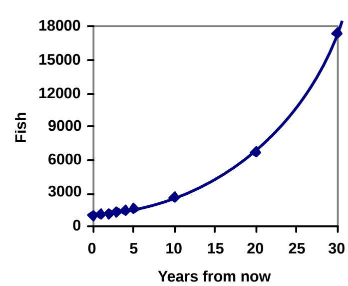
$A = 1000(1.10)^n$

From this, we can quickly calculate the number of fish in 10, 20, or 30 years:

$A = 1000(1.10)^{10} = 2594$

$A = 1000(1.10)^{20} = 6727$

$A = 1000(1.10)^{30} = 17449$



Adding these values to our graph reveals a shape that is definitely not linear. If our fish population had been growing linearly by 100 fish each year, the population would have only reached 4000 in 30 years compared to almost 18000 with this percent-based growth, called exponential growth.

In exponential growth, the population grows proportional to the size of the population, so as the population gets larger, the same percent growth will yield a larger numeric growth.

Exponential Growth

If a quantity starts at size P and grows by $R\%$ (written as a decimal, r) every time period, then the amount A after n years can be determined using the following formula:

$A = P(1+r)^n$

We call r the **growth rate**.

The term $(1+r)$ is called the **growth multiplier**, or common ratio.

Example 1

Between 2007 and 2008, Olympia, WA grew almost 3% to a population of 245 thousand people. If this growth rate was to continue, what would the population of Olympia be in 2014?

Solution

As we did before, we first need to define what year will correspond to $n = 0$. Since we know the population in 2008, it would make sense to have 2008 correspond to $n = 0$, then year 2014 would then be $n = 6$. We have $P = 245,000$.

We know the growth rate is 3%, giving $r = 0.03$.

Using the explicit form:

$A = 245,000(1+0.03)^6 = 245,000(1.19405) = 292,542.25$

The model predicts that in 2014, Olympia would have a population of about 293 thousand people.

Evaluating exponents on the calculator

To evaluate expressions like $(1.03)^6$, it will be easier to use a calculator than multiply 1.03 by itself six times. Most scientific calculators have a button for exponents. It is typically either labeled like:

[^], $[y^x]$, or $[x^y]$.

To evaluate 1.03^6 we'd type 1.03 [^] 6, or 1.03 $[y^x]$ 6. Try it out - you should get an answer around 1.1940523.

Try it Now 1

India is the second most populous country in the world, with a population in 2008 of about 1.14 billion people. The population is growing by about 1.34% each year. If this trend continues, what will India's population grow to by 2020?

Answer

Using $n = 0$ corresponding with 2008,

$A = 1.14(1+0.0134)^{12} =$ about 1.337 billion people in 2020

Example 2

A friend is using the equation $A = 4600(1.072)^n$ to predict the annual tuition at a local college. She says the formula is based on the number of years after 2010. What does this equation tell us?

Solution

In the equation, $P = 4600$, which is the starting value of the tuition when $n = 0$. This tells us that the tuition in 2010 was $4,600.

The growth multiplier is 1.072, so the growth rate is 0.072, or 7.2%. This tells us that the tuition is expected to grow by 7.2% each year.

Putting this together, we could say that the tuition in 2010 was $4,600, and is expected to grow by 7.2% each year.

Evaluating roots on the calculator

In the previous example, we had to calculate the 10th root of a number. This is different than taking the basic square root, $\sqrt{}$. Many scientific calculators have a button for general roots. It is typically labeled like:

[$\sqrt[n]{}$], [$\sqrt[x]{}$], or [$\sqrt[y]{x}$]

To evaluate the 3rd root of 8, for example, we'd either type 3 [$\sqrt[x]{}$] 8, or 8 [$\sqrt[x]{}$] 3, depending on the calculator. Try it on yours to see which to use – you should get an answer of 2.

If your calculator does not have a general root button, all is not lost. You can instead use the property of exponents which states that $\sqrt[n]{a} = a^{1/n}$. So, to compute the 3rd root of 8, you could use your calculator's exponent key to evaluate $(2)^2 - 1 =$. To do this, type:

8 $[y^x]$ (1 [÷] 3)

The parentheses tell the calculator to divide 1/3 before doing the exponent.

Example 3

In 1990, the residential energy use in the US was responsible for 962 million metric tons of carbon dioxide emissions. By the year 2000, that number had risen to 1182 million metric tons[1]. If the emissions **grow exponentially** and continue at the same rate, what will the emissions grow to by 2050?

Solution

Similar to before, we will correspond $n = 0$ with 1990, as that is the year for the first piece of data we have. That will make $P = 962$ (million metric tons of CO_2). In this problem, we are not given the growth rate, but instead are given that $A = 1182$, the amount in 2000 (after 10 years).

When $n = 10$, the explicit equation looks like:

$A = P(1+r)^{10}$

We know the value for P, so we can put that into the equation:

$A = 962(1+r)^{10}$

We also know that $A = 1182$, so substituting that in, we get

$1182 = 962(1+r)^{10}$

We can now solve this equation for the growth rate, r. Start by dividing by 962.

$$\frac{1182}{962} = (1+r)^{10} \qquad \text{Take the 10th root of both sides}$$

$$\sqrt[10]{\frac{1182}{962}} = 1+r \qquad \text{Subtract 1 from both sides}$$

$$r = \sqrt[10]{\frac{1182}{962}} - 1 = 0.0208 = 2.08\%$$

So if the emissions are growing exponentially, they are growing by about 2.08% per year. We can now predict the emissions in 2050, 60 years after 1990, so $n = 60$,

$A = 962(1+0.0208)^{60} = 3308.4$ million metric tons of CO_2 in 2050.

Rounding

As a note on rounding, notice that if we had rounded the growth rate to 2.1%, our calculation for the emissions in 2050 would have been 3347. Rounding to 2% would have changed our result to 3156. A very small difference in the growth rates gets magnified greatly in exponential growth. For this reason, it is recommended to round the growth rate as little as possible.

If you need to round, **keep at least three significant digits** - numbers after any leading zeros. So 0.4162 could be reasonably rounded to 0.416. A growth rate of 0.001027 could be reasonably rounded to 0.00103.

Example 4

Looking back at example 3, for the sake of comparison, what would the carbon emissions be in 2050 if emissions **grow linearly** at the same rate?

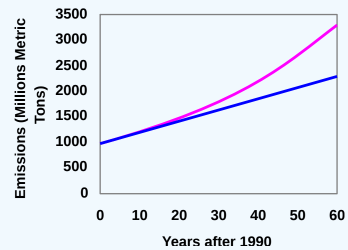
Solution

Again we let $n = 0$ correspond to the year 1990 when the emissions were 962 million metric tons. After 10 years, the emissions increased to 1182 million metric tons. Since we are looking at linear growth, find the slope, then write the equation of the line. This is a good review of section 7.2.

Calculate the slope, m, using the two points (0,962) and (10,1182). We get $m = 22$. Thus, the linear model is:

$y = 22x + 962$

This tells us that if the emissions are changing linearly, they are growing by 22 million metric tons each year. Predicting the emissions in 2050 (60 years after 1990),



$y = 22(60) + 962 = 2282$ million metric tons.

You will notice that this number is substantially smaller than the prediction from the exponential growth model. Calculating and plotting more values helps illustrate the differences.

So how do we know which growth model to use when working with data? There are two approaches which should be used together whenever possible:

1) Find more than two pieces of data. Plot the values, and look for a trend. Does the data appear to be changing like a line, or do the values appear to be curving upwards?

2) Consider the factors contributing to the data. Are they things you would expect to change linearly or exponentially? For example, in the case of carbon emissions, we could expect that, absent other factors, they would be tied closely to population values, which tend to change exponentially.

Try it Now 2

The number of users on a social networking site was 45 thousand in February when they officially went public, and grew to 60 thousand by October. If the site is growing exponentially, and growth continues at the same rate, how many users should they expect two years after they went public?

Answer

Here we will measure n in months rather than years, with $n = 0$ corresponding to February when the site went public. This gives $P = 45$ thousand. October is 8 months later, so $A = 60$ when $n = 8$.

$A = P(1+r)^n$

$60 = 45(1+r)^8$

$\frac{60}{45} = (1+r)^8$

$\sqrt[8]{\frac{60}{45}} = 1 + r$

$r = \sqrt[8]{\frac{60}{45}} - 1 = 0.0366$, or 3.66%

The general explicit equation is $A = 45(1.0366)^n$.

Predicting 24 months (2 years) after they went public:

$A = 45(1.0366)^{24} = 106.63$ thousand users.

Doubling Time Model

Example 5: *E. coli* Bacteria

A water tank up on the San Francisco Peaks is contaminated with a colony of 80,000 *E. coli* bacteria. The population doubles every five days. We want to find a model for the population of bacteria present after t days. The amount of time it takes the population to double is five days, so this is our time unit. After t days have passed, then $t/5$ is the number of time units that have passed. Starting with the initial amount of 80,000 bacteria, our doubling model becomes:

$$A = 80,000(2)^{\frac{t}{5}}$$

Using this model, how large is the colony in two weeks' time? We have to be careful that the units on the times are the same; 2 weeks = 14 days.

Solution

$$A = 80,000(2)^{\frac{14}{5}} = 557,152$$

The colony is now 557,152 bacteria.

Definition: Doubling Time Model

If D is the doubling time of a quantity (the amount of time it takes the quantity to double) and P is the initial amount of the quantity then the amount of the quantity present after t units of time is

$$A = P(2)^{\frac{t}{D}} \tag{7.4.1}$$

Example 6: Files

The doubling time of a population of flies is eight days. If there are initially 100 flies, how many flies will there be in 17 days?

Solution

To solve this problem, use the doubling time model with $D = 8$ and $P = 100$. So the doubling time model for this problem is:

$$A = 100(2)^{t/8}$$

When $t = 17\ days$,

$$A = 100(2)^{\frac{17}{8}} = 436$$

There are 436 flies after 17 days.

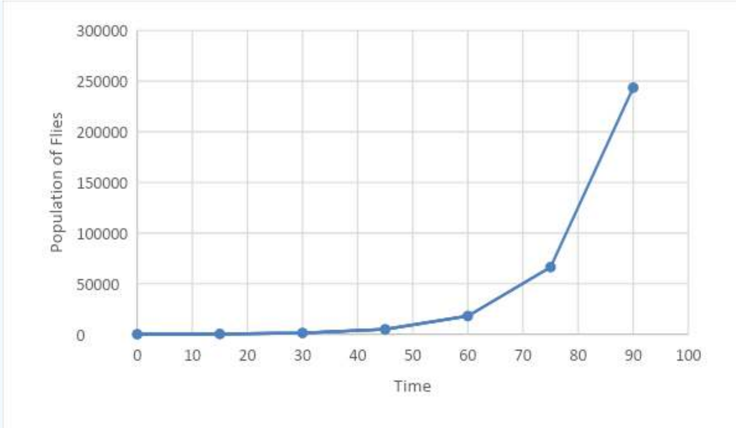


Figure 7.4.6: Graph of Fly Population in Three Months

Note: The population of flies follows an exponential growth model.

Sometimes we want to solve for the length of time it takes for a certain population to grow given their doubling time. To solve for the exponent, we use the *log* button on the calculator.

Example 7: Bacteria Growth

Suppose that a bacteria population doubles every six hours. If the initial population is 4000 individuals, how many hours would it take the population to increase to 25,000?

Solution

We know that $P = 4000$, $D = 6$, and $A = 25000$, so the doubling time model for this problem is:

$$25000 = 4000(2)^{\frac{t}{6}}$$

We need to solve for t

$$\frac{25,000}{4000} = \frac{4000(2)^{\frac{t}{6}}}{4000}$$

$$6.25 = (2)^{\frac{t}{6}}$$

Now, take the log of both sides of the equation

$$\log(6.25) = \log(2)^{\frac{t}{6}}$$

Bring the exponent out front using rules of logarithms

$$\log(6.25) = (\frac{t}{6})\log(2)$$

Divide both sides of the equation by $\log(2)$

$$\frac{\log(6.25)}{\log(2)} = \frac{t}{6}$$

Simplify using your calculator

$$2.6439 = \frac{t}{6}$$

$$t = 15.8631$$

The population would increase to 25,000 bacteria in approximately 15.9 hours.

Example 8: Bird Population

A bird population on a certain island has an annual growth rate of 2.5% per year. The population doubling time is given as 28 years. Approximate the number of years it will take the population to double. If the initial population is 20 birds, use it to find the bird population of the island in 17 years.

Solution

With the bird population doubling in 28 years, we use the doubling time model to find the population is 17 years.

$$A = 20(2)^{\frac{t}{28}}$$

When $t = 17$ years

$$A = 20(2)^{\frac{17}{28}} = 30.46$$

There will be 30 birds on the island in 17 years.

Exponential Decay and Half-Life Model

The half-life of a material is the time it takes for a quantity of material to be cut in half. This term is commonly used when describing radioactive metals like uranium or plutonium. For example, the half-life of carbon-14 is 5730 years.

If a substance has a half-life, this means that half of the substance will be gone in a unit of time. In other words, the amount decreases by 50% per unit of time.

Definition: Half-Life Model

If H is the half-life of a quantity (the amount of time it takes the quantity be cut in half) and P is the initial amount of the quantity then the amount of the quantity present after t units of time is

$$A = P(\frac{1}{2})^{\frac{t}{H}} \tag{7.4.2}$$

Example 9: Half-Life

Let's say a substance has a half-life of eight days. If there are 40 grams present now, how much is left after three days?

Solution

We want to find a model for the quantity of the substance that remains after *t* days. The amount of time it takes the quantity to be reduced by half is eight days, so this is our time unit. After *t* days have passed, then t8 is the number of time units that have passed. Starting with the initial amount of 40, our half-life model becomes:

$$A = 40(\frac{1}{2})^{\frac{t}{8}}$$

With $t = 3$

$$A = 40(\frac{1}{2})^{\frac{3}{8}} = 30.8$$

There are 30.8 grams of the substance remaining after three days.

Example 10: Lead-209

Lead-209 is a radioactive isotope. It has a half-life of 3.3 hours. Suppose that 40 milligrams of this isotope is created in an experiment, how much is left after 14 hours?

Solution

Use the half-life model to solve this problem.

$P = 40$ and $H = 3.3$, so the half-life model for this problem is:

$$A = 40(\frac{1}{2})^{\frac{t}{3.3}}$$

With $t = 14$ hours,

$$A = 40(\frac{1}{2})^{\frac{14}{3.3}} = 2.1$$

There are 2.1 milligrams of Lead-209 remaining after 14 hours.

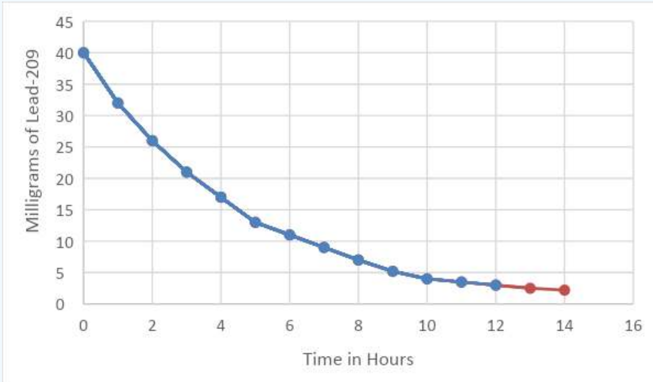


Figure 7.4.10: Lead-209 Decay Graph

Note: The milligrams of Lead-209 remaining follows a decreasing exponential growth model.

Example 11: Carbon-14

Radioactive carbon-14 is used to determine the age of artifacts because it concentrates in organisms only when they are alive. It has a half-life of 5730 years. In 1947, earthenware jars containing what are known as the Dead Sea Scrolls were found. Analysis indicated that the scroll wrappings contained 76% of their original carbon-14. Estimate the age of the Dead Sea Scrolls.

Solution

In this problem, we want to estimate the age of the scrolls, so we need to find t. In 1947, 76% of the carbon-14 remained. This means that the amount, A, remaining at time t, divided by the original amount of carbon-14, P, is equal to 76%.

Thus, $\frac{A}{P} = 0.76$. Use this fact to solve for t.

$$A = P(\frac{1}{2})^{\frac{t}{5730}}$$
$$\frac{A}{P} = (\frac{1}{2})^{\frac{t}{5730}}$$
$$0.76 = (\frac{1}{2})^{\frac{t}{5730}}$$

Now, take the log of both sides of the equation. And let's rewrite the fraction as a decimal.

$$\log(0.76) = \log(0.5)^{\frac{t}{5730}}$$

Bring the exponent out front using rules of logarithms

$$\log(0.76) = (\frac{t}{5730})\log(0.5)$$

Divide both sides of the equation by $\log(0.5)$

$$\frac{\log(0.76)}{\log(0.5)} = \frac{t}{5730}$$

Simplify using your calculator

$$0.3959 = \frac{t}{5730}$$
$$t = 2268.6713$$

The Dead Sea Scrolls are over 2268 years old.

? Try it Now 3

The population of wild elephants is decreasing by 7% per year which gives the half -life for this population to be approximately 10 years. If there are currently 7000 elephants left in the wild, and the population continues to decrease at this rate, how many elephants will remain in the wild in 25 years?

Answer

$P = 7000$ and $H = 10$ years, so the half-life model for this problem is:

$$A = 7000(\frac{1}{2})^{\frac{t}{10}}$$

When $t = 25$,

$$A = 7000(\frac{1}{2})^{\frac{25}{10}} = 1237.4$$

There will be approximately 1237 wild elephants left in 25 years.

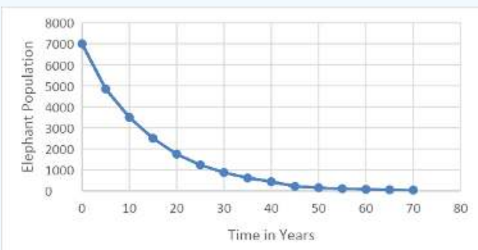


Figure $7.4.3$ Elephant Population over a 70 Year Span.

Note: The population of elephants follows a decreasing exponential growth model.

Review of Exponent Rules and Logarithm Rules

Rules of Exponents	Rules of Logarithm for the Common Logarithm (Base 10)
Definition of an Exponent $a^n = a \cdot a \cdot a \cdot a \cdot \ldots\ldots \cdot a$ (n a's multiplied together)	**Definition of a Logarithm** $10^y = x$ if and only if $\log x = y$
Zero Rule $a^0 = 1$	
Product Rule $a^m \cdot a^n = a^{m+n}$	**Product Rule** $\log(xy) = \log(x) + \log(y)$
Quotient Rule $\frac{a^m}{a^n} = a^{m-n}$	**Quotient Rule** $\log(\frac{x}{y}) = \log(x) - \log(y)$
Power Rule $(a^n)^m = a^{n \cdot m}$	**Power Rule** $\log x^r = r\log(x)(x > 0)$
Distributive Rules $(ab)^n = a^n \cdot a^n, (\frac{a}{b})^n = \frac{a^n}{b^n}$	$\log 10^x = x\log(10) = x$
Negative Exponent Rules $a^{-n} = \frac{1}{a^n}, (\frac{a}{b})^{-n} = (\frac{b}{a})^n$	$10^{\log x} = x(x > 0)$

[1] www.eia.doe.gov/oiaf/1605/ggrpt/carbon.html

CHAPTER OVERVIEW

8: Consumer Mathematics

8.1: Percent

We have to work with money every day. Calculating your monthly expenses, splitting the tab on a restaurant bill, or figuring out how much the tip should be, requires only arithmetic. But when we start saving for the future, planning for retirement, or need a loan, then we need more mathematics.

When one hears the word “percent,” other words come immediately to mind, words such as “century,” “cents,” or “centimeters.” A century equals 100 years. There are one hundred cents in a dollar and there are 100 centimeters in a meter. Thus, it should come as no surprise that percent means “parts per hundred.”

The Meaning of Percent

In the square shown in Figure 8.1.1, a large square has been partitioned into ten rows of ten little squares in each row. We have shaded 20 of 100 possible little squares, or 20% of the total number of little squares.

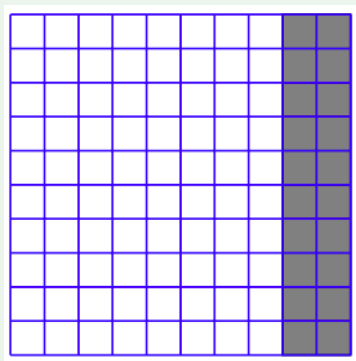


Figure 8.1.1: Shading 20 of 100 little squares, or 20% of the total number of little squares.

Notice in the figure that 80 out of a possible 100 squares are left unshaded. Thus, 80% of the little squares are unshaded. If instead we shaded 35 out of the 100 squares, then 35% of the little squares would be shaded. If we shaded all of the little squares, then 100% of the little squares would be shaded (100 out of 100).

So, when you hear the word “percent,” think “parts per hundred.”

Changing a Percent to a Fraction

Based on the discussion above, it is fairly straightforward to change a percent to a fraction.

Percent to Fraction

To change a percent to a fraction, drop the percent sign and put the number over 100.

Example 1

Change 24% to a fraction.

Solution

Drop the percent symbol and put 24 over 100.

$$\begin{aligned} 24\% &= \frac{24}{100} \quad \color{red}{\text{Percent: Parts per hundred.}} \\ &= \frac{6}{25} \quad \color{red}{\text{Reduce.}} \end{aligned}$$

Hence, 24% = 6/25.

Exercise

Change 36% to a fraction reduced to lowest terms.

Answer

9/25

Example 2

Change $14\frac{2}{7}\%$ to a fraction.

Solution

Drop the percent symbol and put $14\frac{2}{7}$ over 100.

$$\begin{aligned} 14\frac{2}{7}\% &= \frac{14\frac{2}{7}}{100} && \text{Percent: Parts per hundred.} \\ &= \frac{\frac{100}{7}}{100} && \text{Mixed to improper fraction.} \\ &= \frac{100}{7} \cdot \frac{1}{100} && \text{Invert and multiply.} \\ &= \frac{\cancel{100}}{7} \cdot \frac{1}{\cancel{100}} && \text{Cancel.} \\ &= \frac{1}{7} \end{aligned}$$

Hence, $14\frac{2}{7}\% = 1/7$.

Exercise

Change $11\frac{1}{9}\%$ to a fraction reduced to lowest terms.

Answer

1/9

Example 3

Change 28.4% to a fraction.

Solution

Drop the percent symbol and put 28.4 over 100.

$$\begin{aligned} 28.4\% &= \frac{28.4}{100} && \text{Percent: Parts per hundred.} \\ &= \frac{28.4 \cdot 10}{100 \cdot 10} && \text{Multiply numerator and denominator by 10.} \\ &= \frac{284}{1000} && \text{Multiplying by 10 moves decimal point one place right.} \\ &= \frac{71 \cdot 4}{250 \cdot 4} && \text{Factor.} \\ &= \frac{71}{250} && \text{Cancel common factor.} \end{aligned}$$

Exercise

Change 87.5% to a fraction reduced to lowest terms.

Answer

7/8

Changing a Percent to a Decimal

To change a percent to a decimal, we need only remember that percent means "parts per hundred."

Example 4

Change 23.25% to a decimal.

Solution

Drop the percent symbol and put 23.25 over 100.

$$\begin{aligned} 23.25\% &= \frac{23.25}{100} && \text{Percent: Parts per hundred.} \\ &= 0.2325 && \text{Dividing by 100 moves decimal point 2 places left.} \end{aligned}$$

Therefore, 23.25% = 0.2325.

Exercise

Change 2.4% to a decimal.

Answer

0.024

This last example motivates the following simple rule.

Percent to a Decimal

To change a percent to a decimal, drop the percent symbol and move the decimal point two places to the left.

Example 5

Change $5\frac{1}{2}\%$ to a decimal.

Solution

Note that 1/2=0.5, then move the decimal 2 places to the left.

$$\begin{aligned} 5\frac{1}{2}\% &= 5.5\% && 1/2 = 0.5. \\ &= 0.055 && \text{Drop \% symbol.} \\ & && \text{Move decimal point 2 places left.} \\ &= 0.055 \end{aligned}$$

Thus, $5\frac{1}{2}\% = 0.055$.

Exercise 8.1.1

Change $6\frac{3}{4}\%$ to a decimal.

Answer

0.0675

Changing a Decimal to a Percent

Changing a decimal to a percent is the exact opposite of changing a percent to a decimal. In the latter case, we drop the percent symbol and move the decimal point 2 places to the left. The following rule does just the opposite.

Decimal to a Percent

To change a decimal to a percent, move the decimal point two places to the right and add a percent symbol.

Example 6

Change 0.0725 to a percent.

Solution

Move the decimal point two places to the right and add a percent symbol.

$$\begin{aligned} 0.0725 &= 007.25\% \\ &= 7.25\% \end{aligned}$$

Exercise

Change to 0.0375 to a percent.

Answer

3.75%

Example 7

Change 1.025 to a percent.

Solution

Move the decimal point two places to the right and add a percent symbol.

$$\begin{aligned} 1.025 &= 102.5\% \\ &= 102.5\% \end{aligned}$$

Exercise

Change 0.525 to a percent.

Answer

52.5%

Changing a Fraction to a Percent

One way to proceed is to first change the fraction to a decimal, then change the resulting decimal to a percent.

Fractions to Percents: Technique #1

To change a fraction to a percent, follow these steps:

1. Divide numerator by the denominator to change the fraction to a decimal.
2. Move the decimal point in the result two places to the right and append a percent symbol.

Example 8

Use Technique #1 to change 5/8 to a percent.

Solution

Change 5/8 to a decimal, then change the decimal to a percent.

To change 5/8 to a decimal, divide 5 by 8. Since the denominator is a product of twos, the decimal should terminate.

$$\begin{array}{r} 0.625 \\ 8\overline{)5.000} \\ \underline{4\,8} \\ 20 \\ \underline{16} \\ 40 \\ \underline{40} \\ 0 \end{array}$$

To change 0.625 to a percent, move the decimal point 2 places to the right and append a percent symbol.

0.625 = 0 62.5% = 62.5%

Exercise

Change 5/16 to a percent.

Answer

31.35%

A second technique is to create an equivalent fraction with a denominator of 100.

Fractions to Percents: Technique #2

To change a fraction to a percent, create an equivalent fraction with a denominator of 100.

Example 9

Use Technique #2 to change 5/8 to a percent.

Solution

Create an equivalent fraction for 5/8 with a denominator of 100.

$$\frac{5}{8} = \frac{x}{100}$$

Solve this proportion for *x*.

$$\begin{aligned} 8x &= 500 && \color{red}{\text{Cross multiply.}} \\ \frac{8x}{8} &= \frac{500}{8} && \color{red}{\text{Divide both sides by 8.}} \\ x &= \frac{125}{2} && \color{red}{\text{Reduce: Divide numerator and denominator by 4.}} \\ x &= 62.5 && \color{red}{\text{Divide.}} \end{aligned}$$

Thus,

$$\frac{5}{8} = \frac{62.5}{100} = 62.5\%.$$

Alternate Ending

We could also change 125/2 to a mixed fraction; i.e., 125/2 = 62 1 2 . Then,

$$\frac{5}{8} = \frac{62\frac{1}{2}}{100} = 62\frac{1}{2}\%.$$

Same answer.

? Exercise

Change 4/9 to a percent.

Answer

$44\frac{4}{9}\%$

Sometimes we will be content with an approximation.

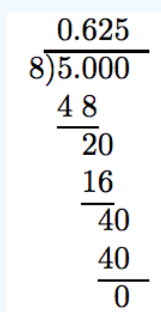
✔ Example 10

Change 4/13 to a percent. Round your answer to the nearest tenth of a percent.

Solution

We will use Technique #1.

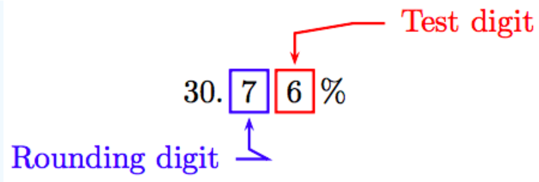
To change 4/13 to a decimal, divide 4 by 13. Since the denominator has factors other than 2’s and 5’s, the decimal will repeat. However, we intend to round to the nearest tenth of a percent, so we will carry the division to four decimal places only. (*Four places are necessary because we will be moving the decimal point two places to the right.*)



To change the decimal to a percent, move the decimal point two places to the right.

0.3076 ≈ 0 30.76% ≈ 30.76%

To round to the nearest tenth of a percent, identify the rounding and test digits.



Because the test digit is greater than or equal to 5, add 1 to the rounding digit and truncate. Thus,

$$0.03076 \approx 30.8\%.$$

Exercise

Change 4/17 to a percent. Round your answer to the nearest tenth of a percent.

Answer

23.5%

8.2: Simple and Compound Interest

Money is not free to borrow! We will refer to money in terms of **present value P,** which is an amount of money at the present time, and **future value F,** which is an amount of money in the future. Usually, if someone loans money to another person in present value, and are promised to be paid back in future value, then the person who loaned the money would like the future value to be more than the present value. That is because the value of money declines over time due to inflation. Therefore, when a person loans money, they will charge interest. They hope that the interest will be enough to beat inflation and make the future value more than the present value.

Simple Interest

Simple interest is interest that is only calculated on the initial amount of the loan (present value, P). This means you are paying the same amount of interest every year. An example of simple interest is when someone purchases a U.S. Treasury Bond.

Simple Interest

To calculate simple interest only,

$$I = Prt \tag{8.2.1}$$

where,

- P is the Present value
- r is the Annual percentage rate (APR) changed to a decimal
- t is the Number of years

Future Value

To calculate the future value based on simple interest,

$$F = P(1 + rt) \tag{8.2.2}$$

where,

- F is the Future value
- P is the Present value
- r is the Annual percentage rate (APR) changed to a decimal
- t is the Number of years

Example 8.2.1: Simple Interest -- Using a Table

Sue borrows $2000 at 5% annual simple interest from her bank. How much does she owe after five years?

Solution

Table 8.2.1: Simple Interest Using a Table

Year	Interest Earned	Total Balance Owed
1	$2000*.05 = $100	$2000 + $100 = $2100
2	$2000*.05 = $100	$2100 + $100 = $2200
3	$2000*.05 = $100	$2200 + $100 = $2300
4	$2000*.05 = $100	$2300 + $100 = $2400
5	$2000*.05 = $100	$2400 + $100 = $2500

After 5 years, Sue owes $2500.

Example 8.2.2: Simple Interest -- Using the Formula

Chad got a student loan for $10,000 at 8% annual simple interest. How much does he owe after one year? How much interest will he pay for that one year?

Solution

P = $10,000, r = 0.08, t = 1

$$F = P(1 + rt) \tag{8.2.3}$$

$$F = 10000(1 + 0.08(1)) = \$10,800 \tag{8.2.4}$$

Chad owes $10,800 after one year. He will pay $10800 - $10000 = $800 in interest.

Example 8.2.3: Simple Interest -- More than 1 year

Carlos deposits $20,000 into a savings account earning 7.25% annual simple interest. How much does he have in the account after 6 years? What was the total interest earned?

Solution

P = $20,000, r = 0.0725, t = 6

$$F = P(1 + rt) \tag{8.2.5}$$

$$F = 20000(1 + 0.0725(6)) = \$28,700 \tag{8.2.6}$$

Carlos has $28,700 in his account after 6 years. He earned $28,700 - $20000 = $8,700 in interest.

Example 8.2.4: Simple Interest -- Finding Time

Ben wants to buy a used car. He has $3000 but wants $3500 to spend. He invests his $3000 into an account earning 6% annual simple interest. How long will he need to leave his money in the account to accumulate the $3500 he wants?

Solution

F = $3500, P = $3000, r = 0.06

$$F = P(1 + rt) \tag{8.2.7}$$

$$\begin{aligned} 3500 &= 3000(1 + 0.06t) \\ \frac{3500}{3000} &= 1 + 0.06t \\ \frac{3500}{3000} - 1 &= 0.06t \\ \frac{\frac{3500}{3000} - 1}{0.06} &= t \end{aligned}$$

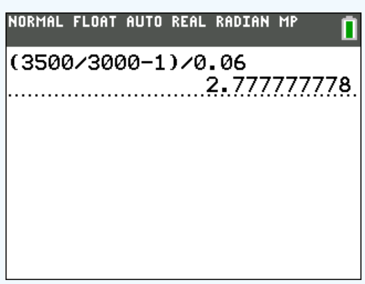


Figure 8.2.2: Calculation to Find t on a TI 83/84 Calculator

$$t \approx 2.8 \text{ years}$$

Ben would need to invest his $3000 for about 2.8 years until he would have $3500 to spend on a used car.

Note: As shown above, wait to round your answer until the very last step so you get the most accurate answer.

Compound Interest

Compound interest is interest paid both on the original principal and on all interest that has been added to the original principal.

Most banks, loans, credit cards, etc. charge you compound interest, not simple interest. This is interest paid on the principal AND the interest accrued. Interest on a mortgage or auto loan is compounded monthly. Interest on a savings account can be compounded quarterly (four times a year). Interest on a credit card can be compounded weekly or daily!

Table 8.2.1: Compounding Periods

Compounding type	Number of compounding periods per year, m
Annually	1
Semiannually	2
Quarterly	4
Monthly	12
Daily	365

Example 8.2.5: Compound Interest -- Using a Table

Suppose you invest $3000 into an account that pays you 7% interest per year for four years. Using compound interest, after the interest is calculated at the end of each year, then that amount is added to the total amount of the investment. Then the following year, the interest is calculated using the new total of the loan.

Table 8.2.2: Compound Interest Using a Table

Year	Interest Earned	Total of Loan
1	$3000*0.07 = $210	$3000 + $210 = $3210
2	$3210 *0.07 = $224.70	$3210 + $224.70 = $3434.70
3	$3434.70*0.07 = $240.43	$3434.70 + $240.43 = $3675.13
4	$3675.13 *0.07 = $257.26	$3675.13 + $257.26 = $3932.39
Total	$932.39	

So, after four years, you have earned $932.39 in interest for a total of $3932.39.

Compound Interest Formula

$$F = P\left(1+\frac{r}{m}\right)^{mt} \tag{8.2.8}$$

where

- F = Future value
- P = Present value
- r = Annual percentage rate (APR) changed into a decimal
- t = Number of years
- m = Number of compounding periods per year

Example 8.2.6: Comparing Simple Interest versus Compound Interest

Let's compare a savings plan that pays 6% simple interest versus another plan that pays 6% annual interest compounded quarterly. If we deposit $8,000 into each savings account, how much money will we have in each account after three years?

Solution

6% Simple Interest: P = $8,000, r = 0.06, t = 3

$$F = P(1+rt) \tag{8.2.9}$$

$$F = 8000(1+0.06(3)) \tag{8.2.10}$$

$$F = 9440 \tag{8.2.11}$$

Thus, we have $9440.00 in the simple interest account after three years.

6% Interest Compounded Quarterly: P = $8,000, r = 0.06, t = 3, m=4

$$F = P\left(1+\frac{r}{m}\right)^{mt} \tag{8.2.12}$$

$$F = 8000\left(1+\frac{0.06}{4}\right)^{4(3)} \tag{8.2.13}$$

$$F = 8000\left(1+\frac{0.06}{4}\right)^{12} \tag{8.2.14}$$

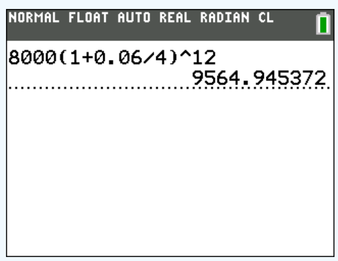


Figure 8.2.3: Calculation for F for Example 8.2.2

$$F = 9564.95 \tag{8.2.15}$$

So, we have $9564.95 in the compounded quarterly account after three years.

With simple interest we earn $1440.00 on our investment, while with compound interest we earn $1564.95.

Example 8.2.7: Compound Interest -- Compounded Monthly

In comparison with Example 8.2.6 consider another account with 6% interest compounded monthly. If we invest $8000 in this account, how much will there be in the account after three years?

Solution

P = $8,000, r = 0.06, t = 3, m = 12

$$F = P\left(1+\frac{r}{m}\right)^{mt} \tag{8.2.16}$$

$$F = 8000\left(1+\frac{0.06}{12}\right)^{12(3)} \tag{8.2.17}$$

$$F = 8000\left(1+\frac{0.06}{12}\right)^{36} \tag{8.2.18}$$

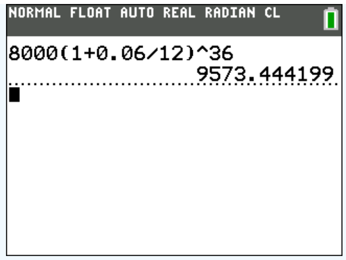


Figure 8.2.4: Calculation for F for Example 8.2.3

$$F = 9573.44 \tag{8.2.19}$$

Thus, we will have $9573.44 in the compounded monthly account after three years.

Interest compounded monthly earns you $9573.44 - $9564.95 = $8.49 more than interest compounded quarterly.

✓ Example 8.2.8: Compound Interest -- Savings Bond

Sophia's grandparents bought her a savings bond for $200 when she was born. The interest rate was 3.28% compounded semiannually, and the bond would mature in 30 years. How much will Sophia's bond be worth when she turns 30?

Solution

P = $200, r = 0.0328, t = 30, m = 2

$$F = P\left(1+\frac{r}{m}\right)^{mt} \tag{8.2.20}$$

$$F = 200\left(1+\frac{0.0328}{2}\right)^{2(30)} \tag{8.2.21}$$

$$F = 8000\left(1+\frac{0.0328}{2}\right)^{60} \tag{8.2.22}$$

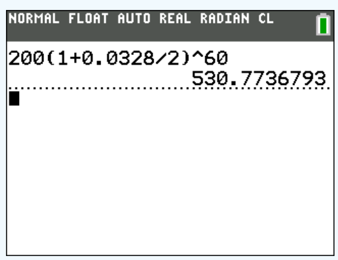


Figure 8.2.5: Calculation for F for Example 8.2.4

$$F = 530.77 \tag{8.2.23}$$

Sophia's savings bond will be worth $530.77 after 30 years.

✓ Example 8.2.9: Find the Present Value

Suppose you know that you will need $40,000 for your child's education in 18 years. If your account earns 4% compounded quarterly, how much would you need to deposit now to reach your goal?

Solution

F = $40,000, r = 0.04, t = 18, m = 4

In this case, we're going to have to set up the equation, and solve for P.

$$40000 = P\left(1+\frac{0.04}{4}\right)^{18(4)} \tag{8.2.24}$$

$$40000 = P(1.01)^{72} \tag{8.2.25}$$

$$40000 = P(2.0471) \tag{8.2.26}$$

$$P = \frac{40000}{2.0471} = \$19539.84 \tag{8.2.27}$$

So you would need to deposit $19,539.84 now to have $40,000 in 18 years.

8.3: Credit Cards

Understanding Credit Cards

Revolving credit is an installment loan that remains open to debt increases and credits. These loans require a regular payment schedule but do not have a fixed amount that is being paid off. The balance of the loan changes as the credit is used for products and services as decided by the account holder.

Credit cards are used by a financial institution to give users access to a loan based on revolving credit. When a credit card is used, the account holder is borrowing money from the institution to make a purchase. The maximum a user can borrow is set by the financial institution based on the credit worthiness of the account holder. Each billing period (typically 25-30 days) the account accrues interest for any remaining balance based on the **average daily balance**. The interest is compounded each billing period, which is why it can be easy to get into debt when using credit cards.

Credit Card Myths

Myth 1: Carrying a credit card balance can increase your credit score.

> FALSE! Having a high balance on your cards can actually decrease your credit score since financial institutions see that you have a lot of debt. It is best to pay off the balance of the card every billing period if possible. If not, then keeping your total balance for all credit cards below 15% of the maximum you can borrow is the second-best idea.

Myth 2: Your income is a factor in your credit score.

> FALSE! No institution knows your income except you, your employer, and the IRS. The only way an institution knows your income is if you disclose it to them.

Myth 3: Credit repair companies can fix your credit.

> FALSE! Credit repair companies can only remove information from a credit report that is incorrect, which any user can do for free. Credit repair companies often will create a new loan to consolidate your debt into one payment. This does not necessarily improve the credit score.

Myth 4: Closing old credit cards can hurt your credit score.

> FALSE/TRUE This one is a bit of both.
>
> FALSE: If you have old credit cards that have high maximums that you do not use or excessive number of credit cards, it can be beneficial to close those since financial institutions look at how much debt you COULD get into if you maxed out all your cards.
>
> TRUE: If you have old cards that are not being used, they will show on the credit report as an account that is in good standing which can help your credit score. Some financial institutions will close accounts that have not been used for an extended period of time.

Myth 5: Checking your credit score can hurt your credit score.

> FALSE! When you check your personal credit score it is considered a soft inquiry. These inquiries do not affect the credit score. Everyone can request a free credit report from the three main credit reporting agencies once a year.
>
> Applying for a loan or a new service that requires a credit check is considered a hard inquiry. These can negatively affect your score if there are several hard inquiries within a short period of time for different types of loans. For example: if you are applying to purchase a car at several financial institutions to find the best rate that may not affect your score negatively. If you apply for several different credit cards in a short period of time that will negatively affect your score.

Average Daily Balance

Since the balance of a revolving loan changes often, they typical way of determining interest cannot be used. A financial institution cannot charge interest for the entire billing period on money that was not borrowed at the beginning of that period. For this reason, financial institutions use the average balance over the length of the billing period.

$$\text{average daily balance} = \frac{\text{sum of unpaid balances for each day in the billing period}}{\text{number of days in the billing period}} \tag{8.3.1}$$

✓ Example 1 -- Find the Average Daily Balance

Victor got a new credit card. He made the following charges in the first billing cycle. Determine the average daily balance for his first billing period.

March 1	The first billing cycle begins.
March 3	He bought a jacket for $55.
March 8	He went to the movies and spent $42.
March 15	He put gas in his car and spent $67.
March 21	He treated his family to dinner and spent $73.
March 25	The billing cycle ended.

Solution

To determine the average daily balance we first need to find the sum of unpaid balances for each day in the billing period. The calendar below lists the balances for each day of the billing cycle.

March 1 - $ 0.00	March 2 - $ 0.00	March 3 - $ 55.00	March 4 - $ 55.00	March 5 - $ 55.00
March 6 - $ 55.00	March 7 - $ 55.00	March 8 - $ 97.00	March 9 - $ 97.00	March 10 - $ 97.00
March 11 - $ 97.00	March 12 - $ 97.00	March 13 - $ 97.00	March 14 - $ 97.00	March 15 - $ 164.00
March 16 - $ 164.00	March 17 - $ 164.00	March 18 - $ 164.00	March 19 - $ 164.00	March 20 - $ 164.00
March 21 - $ 237.00	March 22 - $ 237.00	March 23 - $ 237.00	March 24 - $ 237.00	March 25 - $ 237.00

Begin on March 3rd since that is the first date with a balance to find the sum. The sum will be the numerator of the average daily balance fraction.

$$55+55+55+55+55+97+97+97+97+97+97+97+164+164+164+164+164+164+237+237 +237+237+237=3123$$

The denominator will be the number of days in the billing cycle which, in this case, is 25 days.

$$\text{average daily balance} = \frac{3123}{25} = 125.92$$

Interest will be calculated on the amount $125.92 since that is the average daily balance for this billing cycle.

✓ Example 2 -- Calculate the Interest

Jasmine has had her credit card for several months. She only pays the minimum payment, which is 2% of the balance or $10, whichever is more. Review the transactions of her latest billing cycle to determine the average daily balance. Then find the interest that will be charged to her account with a 7.99% APR and the minimum payment required.

June 1	The billing cycle begins with a balance of $2750
June 5	She makes a payment of $55.
June 6	She got a coffee and snacks $27.
June 12	She put gas in her car and spent $33.
June 17	She bought tickets to a concert $119.
June 25	The billing cycle ended.

Solution

In this example, there is an existing balance and a payment is made. The payment will decrease the balance while purchases will increase the balance. Instead of using a calendar to determine the sum, we will use multiplication.

First determine when the balance changed and what the change was. Then determine how many days the account was at the balance.

June 1	$\$2750.00$
June 5	$\$2750 - \$55 = \$2695$
June 6	$\$2695 + \$27 = \$2722$
June 12	$\$2722 + \$33 = \$2755$
June 17	$\$2755 + \$119 = \$2874$
June 30	The billing cycle ended.

4 days	$\$2750$
1 day	$\$2695$
6 days	$\$2722$
5 days	$\$2755$
14 days	$\$2874$

This billing cycle has 30 days. The total number of days should add to 30. Now we can use multiplication to determine the sum of the daily balances.

$$\text{average daily balance} = \frac{4 \times 2750 + 1 \times 2695 + 6 \times 2722 + 5 \times 2755 + 14 \times 2874}{30} = \frac{84038}{30} = 2801.27$$

To find the interest, use the $I = Prt$ formula, where P is the average daily balance, r is the APR, and t is the time in years. Since there are 30 days in the month of June (and 365 in a year) $t = \frac{30}{365}$.

$$I = 2801.27 \times 0.0799 \times \frac{30}{365} = 18.40$$

The interest is added to the balance before the minimum payment is determined. Recall that the minimum payment is either 2% or $10, whichever is greater. The balance is the last value before the billing cycle ended. Do not use the average daily balance to calculate the minimum payment. Note: If the balance is paid in full by the end of the cycle, no interest is charged.

$$2874 + 18.65 = 2892.65 \times 0.02 = 57.85$$

The minimum payment due for this billing period is $57.85.

Try it Now 1

Refer to the following credit card transactions. Determine the average daily balance and the interest charged for this billing cycle for an account with a 24.99% APR. Find the final balance.

May 1	The billing cycle begins with a balance of $4212
May 4	A payment of $25 is made.
May 12	A purchase of $148 is made.
May 16	A purchase of $16 is made.
May 17	A purchase of $96 is made.
May 31	The billing cycle ended.

Answer

Find the sum of the daily balances and divide by the number of days in the billing cycle.

$$\text{average daily balance} = \frac{3 \times 4212 + 8 \times 4187 + 4 \times 4335 + 1 \times 4351 + 15 \times 4447}{31} = \frac{134528}{31} = 4339.61$$

Determine the interest for the current billing cycle.

$$I = 4339.61 \times 0.2499 \times \frac{31}{365} = 90.26$$

The final balance would be $4447 + 90.26 = 4537.26$

Buy Now, Pay Later: An Alternative to Credit Cards?

The payment method known as "Buy Now, Pay Later," or BNPL, has been around for many years, but has become popular recently due to the pandemic. As more consumers buy goods online, merchants have started offering BNPL options for making purchases. You can determine whether a company offers BNPL on the checkout page - you might be given the option to pay off the whole bill immediately or to break it up into regular installments, such as one payment per month. These payments are calculated using the **Add-On Interest Method**.

Add-On Interest

Companies that charge interest for the use of a BNPL plan typically calculate interest on the entire amount borrowed, then add it to the principal. This total is then divided into equal-sized payments.

$$\text{size of payment} = \frac{\text{sum of amount borrowed and interest}}{\text{number of payments}} \tag{8.3.2}$$

More specifically, if the company uses a monthly payment plan, we can use the following formula.

$$\text{monthly payment} = \frac{P+I}{n}$$

where P is the principal, $I = Prt$ is the interest, r is the annual interest rate as a decimal, t is the length of time (in years) before the bill is fully paid off, and n is the number of monthly payments.

Example 3 -- Add-On Interest Method

Jacques is looking to buy a $2,000 couch online. The merchant selling the couch offers a BNPL option for purchasing, where the customer would pay 15% simple interest, and the bill would be spread out in four equal monthly payments. Help Jacques calculate the amount of each monthly payment.

Solution

Since the couch costs $2,000, the principal P is 2,000. We can convert the interest rate to a decimal by moving the decimal point two places to get $r = 0.15$. Since it will take four months (or 1/3 of a year) to pay off the bill, $n = 4$ and $t = 1/3$.

$$I = 2000 \times 0.15 \times \frac{1}{3} = 100$$

and

$$\text{monthly payment} = \frac{2,000 + 100}{4} = 525$$

So Jacques would pay $525 in each of his monthly payments.

When would it be preferable to pay using Buy Now, Pay Later? In general, it is a good rule of thumb to avoid as much debt as possible, especially when for a non-essential item. But there are some situations when it might be better to use BNPL instead of, say, a credit card. You typically do not need to have a credit history to use BNPL, and many payment options charge zero interest. But be careful! These payment plans frequently charge late fees if you do not complete a payment on time, and it is easy to overspend.

8.4: Annuities

Sometimes it makes better financial sense to put small amounts of money away (to save for a period of time time), in order to purchase a large item in the future instead of taking out a loan now with a high interest rate.

Annuity

In the previous sections of this chapter, we examined problems where an amount of money was deposited lump sum in an account and was left there for the entire time period. Now we will solve problems where a series of payments of the same size are made into an account. When a series of payments of some fixed amount are made into an account at equal intervals of time, we call that an **annuity**. When payments are made at the end of each period rather than at the beginning, we call it an **ordinary annuity.**

Future Value of an Ordinary Annuity Formula

The Future Value, F, of an ordinary annuity is the amount in the account, including interest, after making all payments.

$$F = PMT \cdot \frac{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}{\left(\frac{r}{m}\right)} \tag{8.4.1}$$

where,
F = Future value
PMT = Periodic payment/deposit
r = Annual percentage rate (APR) changed to a decimal
t = Number of years
m = Number of compounding periods per year

Example 8.4.1

Tanya deposits \$300 at the end of each quarter in her savings account. If the account earns 5.75% compounded quarterly, how much money will she have in 4 years?

Solution

PMT =300, r =0.0575, t =4, m =4

The future value of this annuity can be found using the above formula.

$$F = 300\frac{\left[\left(1+\frac{0.0575}{4}\right)^{4(4)} - 1\right]}{\left(\frac{0.0575}{4}\right)}$$

$$F = 300\frac{\left[\left(1+\frac{0.0575}{4}\right)^{16} - 1\right]}{\left(\frac{0.0575}{4}\right)}$$

$F = 300(17.8463)$

$F = 5353.89$

If Tanya deposits \$300 into a savings account earning 5.75% compounded quarterly for 4 years, then at the end of 4 years she will have \$5,353.89.

Example 8.4.2

A traditional individual retirement account (IRA) is a special type of retirement account in which the money you invest is exempt from income taxes until you withdraw it. If you deposit $100 each month into an IRA earning 6% interest, how much will you have in the account after 20 years?

Solution

$$\begin{array}{ll} PMT=100 & \text{the monthly deposit} \\ r=0.06 & 6\% \text{ annual rate} \\ t=20 & \text{we want the amount after 20 years} \\ m=12 & \text{since we're doing monthly deposits, we'll compound monthly} \end{array}$$

$$F=100\frac{\left[\left(1+\frac{0.06}{12}\right)^{20(12)}-1\right]}{\left(\frac{0.06}{12}\right)}$$

$$F=100\frac{\left((1.005)^{240}-1\right)}{(0.005)}$$

$$F=100\frac{(3.3102-1)}{(0.005)}$$

$$F=100\frac{(2.3102)}{(0.005)}=\$46204$$

The account will grow to $46,204 after 20 years.

Notice that you deposited into the account a total of $24,000 ($100 a month for 240 months). The difference between what you end up with and how much you put in is the interest earned. In this case it is $\$46,204-\$24,000=\$22,204$

Try it Now

A more conservative investment account pays 3% interest. If you deposit $5 a day into this account, how much will you have after 10 years? How much is from interest?

Answer

$$\begin{array}{ll} PMT=5 & \text{the daily deposit} \\ r=0.03 & 3\% \text{ annual rate} \\ t=10 & \text{we want the amount after 10 years} \\ m=365 & \text{since we're doing daily deposits, we'll compound daily} \end{array}$$

$$F=5\frac{\left[\left(1+\frac{0.03}{365}\right)^{365(10)}-1\right]}{\frac{0.03}{365}}=\$21,282.07$$

We would have deposited a total of $\$5\cdot365\cdot10=\$18,250$ so $3,032.07 is from interest.

Sinking Fund

You may want to save regularly to have a fixed amount available in the future. The account that you establish for your deposits is called a **sinking fund**. Calculating the sinking fund deposit uses the same method as the previous problem. Observe that we are using the words deposit and payment synonymously.

Example 8.4.3

Robert needs $5,000 in three years. How much should he deposit each month in an account that pays 8% compounded monthly in order to achieve his goal?

Solution

We need to find how much Robert should deposit, PMT, each month.

$F = 5000$ the future value
$r = 0.08$ 8% annual rate
$t = 3$ we want the amount after 3 years
$m = 12$ compounded monthly

$$5000 = PMT \frac{\left[\left(1+\frac{0.08}{12}\right)^{12(3)} - 1\right]}{\left(\frac{0.08}{12}\right)}$$

Therefore,

$$\begin{aligned} PMT \frac{\left[\left(1+\frac{.08}{12}\right)^{36)} - 1\right]}{\frac{.08}{12}} &= 5000 \\ \text{PMT}(40.5356) &= 5000 \\ \text{PMT} &= \frac{5000}{40.5356} \\ &= 123.3483 \end{aligned}$$

Robert needs to deposit $123.35 at the end of each month for 3 years into an account paying 8% compounded monthly in order to have $5,000 at the end of 5 years.

Solve the Annuity Formula for Payment, PMT

In the example above, notice we were given the future value and asked to find the payment, *PMT*. It would be helpful to solve the annuity formula for the payment first.

$$F = PMT \cdot \frac{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}{\left(\frac{r}{m}\right)}$$

---Since the *PMT* is multiplied by a fraction, solve for *PMT* by multiplying both sides of the equation by the reciprocal of that fraction.

$$F \cdot \frac{\left(\frac{r}{m}\right)}{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]} = PMT \frac{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}{\left(\frac{r}{m}\right)} \cdot \frac{\left(\frac{r}{m}\right)}{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}$$

---Simply the right side of the equation, then rewrite as,

$$PMT = F \frac{\left(\frac{r}{m}\right)}{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}$$

Sinking Fund Formula

Use this formula when you know the future value and you want to find the payment, *PMT*.

$$PMT = F \cdot \frac{\left(\frac{r}{m}\right)}{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}$$

where,
F = Future value
PMT = Periodic payment/deposit
r = Annual percentage rate (APR) changed to a decimal
t = Number of years
m = Number of compounding periods per year

*Notes:
The compounding frequency is not always explicitly given, but is determined by how often you make payments.
In making payments into a sinking fund, we will always round the payment *up* to the next cent.

Example 8.4.4

You want to have $200,000 in your account when you retire in 30 years. Your retirement account earns 8% interest. How much do you need to deposit each month to meet your retirement goal?

Solution

We need to find the amount of the monthly deposit, PMT.

$$\begin{array}{ll} F = 200,000 & \text{The amount you want to have in 30 years} \\ r = 0.08 & 8\% \text{ annual rate} \\ t = 30 & 30 \text{ years} \\ m = 12 & \text{compounded monthly} \end{array}$$

$$\begin{aligned} PMT &= 200,000 \frac{\left(\frac{0.08}{12}\right)}{\left[\left(1+\frac{0.08}{12}\right)^{30(12)} - 1\right]} \\ PMT &= 200,000(6.7043 \times 10^{-4}) \\ PMT &= 134.086 \end{aligned}$$

You would need to deposit $\$134.09$ each month to have $\$200,000$ in 30 years if your account earns 8% interest.

Example 8.4.5

A business needs $450,000 in five years. How much should be deposited each quarter in a sinking fund that earns 9% compounded quarterly to have this amount in five years?

Solution

Again, suppose that PMT dollars are deposited each quarter in the sinking fund. After five years, the future value of the fund should be $450,000. This suggests the following relationship:

$$PMT = 450,000 \frac{\left(\frac{0.09}{4}\right)}{\left[\left(1+\frac{0.09}{4}\right)^{4(5)} - 1\right]}$$
$$PMT = 450,000(0.040142)$$
$$PMT = 18063.9$$

The business needs to deposit $18,063.90 at the end of each quarter for 5 years into a sinking fund earning 9% interest compounded quarterly in order to have $450,000 at the end of 5 years.

Solving for time, t

Example 8.4.6

Sara wants to save $10,000 for a down payment on a new car. If she deposits $300 each month into an account earning 6.8% interest, how long will it take her to save up the $10,000 she needs?

F = $10,000, PMT = $300, r = 0.068, m = 12

Note: We will use the annuity formula and solve for t.

$$F = PMT \cdot \frac{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}{\left(\frac{r}{m}\right)}$$

$$10000 = 300 \frac{\left[\left(1+\frac{0.068}{12}\right)^{12t} - 1\right]}{\left(\frac{0.068}{12}\right)}$$

$$33.333333 = \frac{(1.005667)^{12t} - 1}{0.005667}$$

$$1.188900 = (1.005667)^{12t}$$

To solve for t, time, take the logarithm (log) of both sides. Recall the “Power Rule” of logs from section 7.4 which allows us to bring the exponent out front.

$$\log x^r = r \log x \tag{8.4.2}$$

$$\log(1.188900) = 12t \cdot \log(1.005667)$$

$$t = \frac{log(1.188900)}{12t \cdot \log(1.005667)}$$

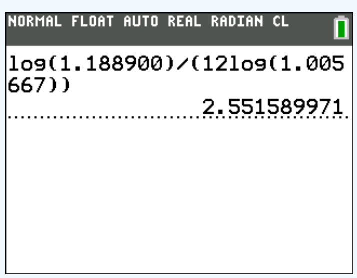


Figure 8.4.6: Calculation for t for Example 8.4.6

It will take Sara about 2.6 years to save up the $10,000.

✓ Example 8.4.7

At the end of each quarter a 50-year old woman puts $1200 in a retirement account that pays 7% interest compounded quarterly. When she reaches age 60, she withdraws the entire amount and places it into a mutual fund that pays 9% interest compounded monthly. From then on she deposits $300 in the mutual fund at the end of each month. How much is in the account when she reaches age 65?

First, calculate the future value for the first 10 years.
$PMT = \$1200$, $r = 0.07$, $t = 10$, $m = 4$

$$F = PMT \cdot \frac{\left[\left(1+\frac{r}{m}\right)^{mt} - 1\right]}{\left(\frac{r}{m}\right)}$$

$$F = 1200\frac{\left[\left(1+\frac{0.07}{4}\right)^{4(10)} - 1\right]}{\left(\frac{0.07}{4}\right)}$$

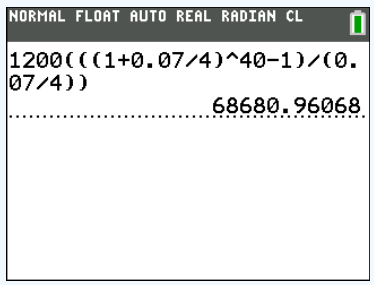


Figure 8.4.7: Calculation for F for Part One of Example 8.4.7

Second, she puts this lump sum plus $300 a month for 5 years at 9%. Think of the lump sum and the new monthly deposits as separate things. The lump sum just sits there earning interest so use the compound interest formula. The monthly payments are a new payment plan, so use the savings plan formula again.

Total = (lump sum + interest) + (new deposits + interest)

$$68,680.96\left(1+\frac{0.09}{12}\right)^{12\cdot 5} + 300\left[\frac{\left(1-\frac{0.09}{12}\right)^{12\cdot 5} - 1}{\frac{0.09}{12}}\right]$$

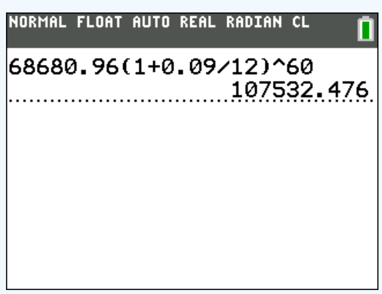


Figure 8.4.8: Calculation for the Lump Sum Plus Interest

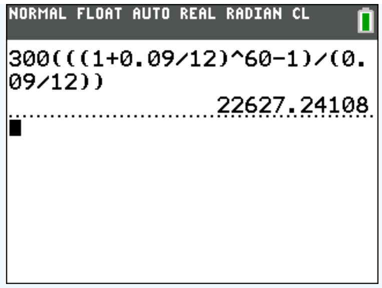


Figure 8.4.9: Calculation for the New Deposits Plus Interest

Total = 107,532.48 + 22,627.24 = 130,159.72

She will have $130,159.72 when she reaches age 65.

8.5: Amortized Loans

In Section 8.2, we learned how to find the future value of a lump sum, and in Section 8.4, we learned how to find the future value of an annuity. With these two concepts in hand, we will now learn how to **amortize a loan**, and how to find the **present value of an annuity**.

Amortize a Loan

If a person or business needs to buy or pay for something now (a car, a home, college tuition, equipment for a business) but does not have the money, they can borrow the money as a loan.

They receive the loan amount now, called the principal, P, (or present value), and are obligated to pay back the principal in the future over a stated amount of time (term of the loan), as regular periodic payments, PMT, plus interest.

Example 8.5.1

Find the monthly payment for a car costing $15,000 if the loan is amortized over five years at an interest rate of 9%.

Solution

Consider the following scenario:

Two people, Mr. Cash and Mr. Credit, go to buy the same car that costs $15,000. Mr. Cash pays cash and drives away, but Mr. Credit wants to make monthly payments for five years.

Our job is to determine the amount that Mr. Credit needs to pay each month for 5 years. We reason as follows:

If Mr. Credit pays PMT dollars per month, then the PMT dollar payment deposited each month at 9% for 5 years should yield the same amount as the $15,000 lump sum deposited in an annuity for 5 years.

Again, we are comparing the future values for both Mr. Cash and Mr. Credit, and we would like them to be the same.

Since Mr. Cash is paying a lump sum of $15,000, the future value F is given by the lump sum formula (8.2), and it is

$$F = 15,000\left(1+\frac{.09}{12}\right)^{60}$$

Mr. Credit wishes to make a sequence of payments of PMT dollars per month, and the future value is given by the annuity formula (8.4), and this value is

$$F = PMT\frac{\left[\left(1+\frac{.09}{12}\right)^{60}-1\right]}{\frac{.09}{12}}$$

We set the two future amounts equal to eachother and solve for the unknown value, PMT.

$$\begin{aligned}15,000\left(1+\frac{.09}{12}\right)^{60} &= PMT\frac{\left[\left(1+\frac{.09}{12}\right)^{60}-1\right]}{\frac{.09}{12}}\\ 15,000(1.5657) &= PMT(75.4241)\\ 311.3792 &= PMT\end{aligned}$$

Therefore, the monthly payment needed to repay the loan is $311.38 for five years.

The formula used above (and restated here), for finding payments on an amortized loan, can appear cumbersome.

$$P\left(1+\frac{r}{m}\right)^{mt} = PMT\frac{\left[\left(1+\frac{r}{m}\right)^{mt}-1\right]}{\frac{r}{m}}$$

If we do the necessary algebra to solve this equation for *PMT*, we can use the new formula to find the payments. The algebra has been omitted and the new formula is stated in the box below.

Amortization Formula

$$PMT = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]}$$

P is the balance in the account at the beginning (the principal, or amount of the loan)

r is the annual interest rate in decimal form

t is the length of the loan, in years

m is the number of compounding periods in one year

PMT is the loan payment (monthly payment, annual payment, etc.)

*Notes:
The compounding frequency is not always explicitly given, but is determined by how often you make payments.
We will round payments on a loan *up* to the next cent.

Example 8.5.2

You want to take out a $340,000 mortgage (home loan). The interest rate on the loan is 3.5%, and the loan is for 30 years. How much will your monthly payments be? How much interest will you pay over the life of the loan?

Solution

We're looking for PMT.

$P = \$340,000$	the starting loan amount
$r = 0.035$	3.5% annual rate
$t = 30$	since we're making monthly payments for 30 years
$m = 12$	since we're doing monthly payments, we'll compound monthly

$$PMT = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]}$$

$$PMT = 340000 \cdot \frac{\left(\frac{0.035}{12}\right)}{\left[1-\left(1+\frac{0.035}{12}\right)^{-12(30)}\right]}$$

$$PMT = 1526.7519$$

You will make payments of $1526.76 per month for 30 years. (Remember to round up for payments.)

You will pay a total of $\$1526.76$per month for 360 months which equals $\$549,633.60$to the loan company.
The total paid over the life of the loan is $\$549,633.60 - \$340,000 = \$209,633.60$

Example 8.5.3

Jack goes to a car dealer to buy a new car for $18,000 at 2% APR with a five-year loan. The dealer quotes him a monthly payment of $425. Verify that this is the correct monthly payment.

P = $18,000, r = 0.02, t = 5, m = 12

$$PMT = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]}$$

$$PMT = 18000 \cdot \frac{\left(\frac{0.02}{12}\right)}{\left[1-\left(1+\frac{0.02}{12}\right)^{-12(5)}\right]}$$

$$PMT = 315.4996$$

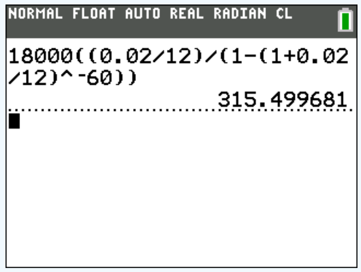


Figure 8.5.3: Calculation for PMT

Jack should have a monthly payment of $315.50, not $425.

What should the total principal and interest be with the $315.50 monthly payment? $\$315.50(12)(5) = \$18,930$
The $315.50 payment per month has a total of $930 in interest paid over the 2-year loan period.

However, the dealer is trying to get Jack to pay $425 per month. This equates to $\$425(12)(5) = \$25,500$ which is significantly more than the calculation above. And notice the difference in the the total interest charges; $25,500 - $18,000 = $7,500. This means that the quoted rate of 2% APR is not accurate, or the quoted price of $18,000 is not accurate, or both.

Try it Now 1

Janine bought $3,000 of new furniture on credit. Because her credit score isn't very good, the store is charging her a fairly high interest rate on the loan: 16%. If she agreed to pay off the furniture over 2 years, how much will she have to pay each month? What is the total interest charged?

Answer

$$\begin{array}{ll} P = 3,000 & \text{the starting loan amount \$3,000 loan} \\ r = 0.16 & 16\% \text{ annual rate} \\ t = 2 & \text{2 year to repay} \\ m = 12 & \text{since we're doing monthly payments, we'll compound monthly} \end{array}$$

$$PMT = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]}$$

$$PMT = 3000 \cdot \frac{\left(\frac{0.16}{12}\right)}{\left[1-\left(1+\frac{0.16}{12}\right)^{-12(2)}\right]}$$

$$PMT = 146.8893$$

Janine will need to make monthly payments of $146.89

In total, she will pay $\$3,525.36$ to the store, meaning she will pay $\$525.36$ in interest over the two years.

Example 8.5.4

With a fixed rate mortgage, you are guaranteed that the interest rate will not change over the life of the loan. Suppose you need $250,000 to buy a new home. The mortgage company offers you two choices: a 30-year loan with an APR of 6% or a 15-year loan with an APR of 5.5%. Compare your monthly payments and total loan cost to decide which loan you should take. Assume no difference in closing costs.

Option 1: First calculate the monthly payment:

$$PMT = P\bullet\left[\frac{\frac{r}{n}}{1-\left(1+\frac{r}{n}\right)^{-nt}}\right]$$

$$PMT = 250000\bullet\left[\frac{\frac{0.06}{12}}{1-\left(1+\frac{0.06}{12}\right)^{-12\cdot 30}}\right]$$

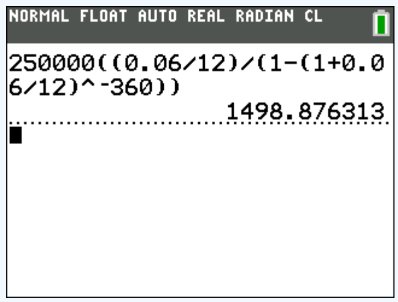


Figure 8.5.5: Calculation for PMT for Example 8.5.5, Option 1

The monthly payment for a 30-year loan at 6% interest is $1498.88.

Now calculate the total cost of the loan over the 30 years:

$$\$1498.88\times 12\times 30=\$539,596.80$$

The monthly payments are $1498.88 and the total cost of the loan is $539,596.80.

Option 2: First calculate the monthly payment:

$$PMT = P\bullet\left[\frac{\frac{r}{n}}{1-\left(1+\frac{r}{n}\right)^{-nt}}\right]$$

$$PMT = 250000\bullet\left[\frac{\frac{0.055}{12}}{1-\left(1+\frac{0.055}{12}\right)^{-12\cdot 15}}\right]$$

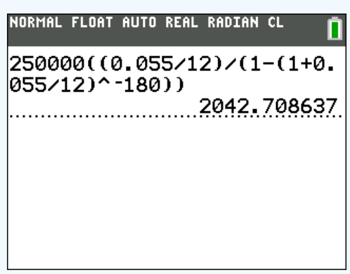


Figure 8.5.6: Calculate PMT for Example 8.5.5, Option 2

The monthly payment for a 15-year loan at 5.5% interest is $2042.71.

Now calculate the total cost of the loan over the 15 years:

$\$2042.71 \times 12 \times 15 = \$367,687.80$

The monthly payments are \$2042.71 and the total cost of the loan is \$367,687.80.

Therefore, the monthly payments are higher with the 15-year loan, but you spend a lot less money overall.

Present Value of an Annuity

The **present value of an annuity** is the amount of money we would need now in order to be able to make the annuity payments in the future. Often, we know how much we can afford to pay for each regular payment, so we need to find how much money we can borrow.

Example 8.5.5

Jordan can afford \$400 per month as a car payment. The car dealership offers an auto loan at 12% interest for 4 years. What is the present value of the car? In other words, what loan amount can Jordan afford at \$400 per month?

Solution

In this example,

$PMT = \$400$ the monthly loan payment
$r = 0.12$ 12% annual rate
$t = 4$ since we're making monthly payments for 4 years
$m = 12$ since we're doing monthly payments, we'll compound monthly

$$PMT = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]}$$

$$400 = P \cdot \frac{\left(\frac{0.12}{12}\right)}{\left[1-\left(1+\frac{0.12}{12}\right)^{-12(4)}\right]}$$

$$P = 15189.5838$$

Jordan will pay a total of \$15,189.58 (\$400 per month for 48 months) to the car dealership. The difference between the amount paid and the amount of the loan is the interest paid. In this case, Jordan is paying $\$19,200 - \$15,189.58 = \$4,010.42$ interest total.

Solve the Amortization Formula for P, Present Value

In the example above, notice we were given the monthly payment and asked to find the the present value, *P*. It would be helpful to solve the amortization formula for the present value first.

$$PMT = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]}$$

---Since *P* is multiplied by a fraction, solve for *P* by multiplying both sides of the equation by the reciprocal of that fraction.

$$PMT \cdot \frac{\left[1-\left(1+\frac{r}{m}\right)^{-m(t)}\right]}{\left(\frac{r}{m}\right)} = P \cdot \frac{\left(\frac{r}{m}\right)}{\left[1-\left(1+\frac{r}{m}\right)^{-mt}\right]} \cdot \frac{\left[1-\left(1+\frac{r}{m}\right)^{-m(t)}\right]}{\left(\frac{r}{m}\right)}$$

---Simply the right side of the equation, then rewrite as,

$$P = PMT \cdot \frac{\left[1-\left(1+\frac{r}{m}\right)^{-m(t)}\right]}{\left(\frac{r}{m}\right)}$$

Present Value of an Annuity Formula

Use this formula when you know the payment and you want to find the present value, *P*.

$$P = PMT \cdot \frac{\left[1-\left(1+\frac{r}{m}\right)^{-m(t)}\right]}{\left(\frac{r}{m}\right)}$$

P is the balance in the account at the beginning (the principal, or amount of the loan)

r is the annual interest rate in decimal form

t is the length of the loan, in years

m is the number of compounding periods in one year

PMT is the loan payment (monthly payment, annual payment, etc.)

*Note: The compounding frequency is not always explicitly given, but is determined by how often you make payments.

Example 8.5.6

Grace buys an iPad from a rent-to-own business on credit with payments of $30 a month for four years at 14.5% APR compounded monthly. If Grace had bought the iPad from Best Buy or Amazon it would have cost $500. What is the price that Grace paid for the iPad at the rent-to-own business? How much interest was paid over the life of the loan? What is the better option?

PMT = $30, *r* = 0.145, *t* = 4, m = 12

$$P = PMT \cdot \frac{\left[1-\left(1+\frac{r}{m}\right)^{-m(t)}\right]}{\left(\frac{r}{m}\right)}$$

$$P = 30 \cdot \frac{\left[1-\left(1+\frac{0.145}{12}\right)^{-12(4)}\right]}{\left(\frac{0.145}{12}\right)}$$

$$P = 1087.8254$$

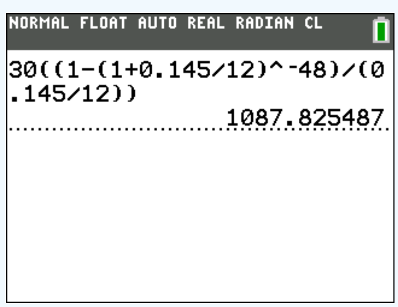


Figure 8.5.5: Calculation of Present Value

The price Grace paid for the iPad was $1,087.83.

That’s a lot more that $500!

Also, the total amount paid over the course of the loan was $\$30 \times 12 \times 4 = \$1440$. Therefore, the total amount of interest paid was $1440 - $1087.83 = $352.17.

Example 8.5.7

Suppose you have won a lottery that pays $1,000 per month for the next 20 years. But, you prefer to have the entire amount now. If the interest rate is 8%, how much will you accept?

Solution

This classic present value problem needs our complete attention because the rationalization we use to solve this problem will be used again in the problems to follow.

Consider, for argument purposes, that two people Mr. Cash, and Mr. Credit have won the same lottery of $1,000 per month for the next 20 years. Mr. Credit is happy with his $1,000 monthly payment, but Mr. Cash wants to have the entire amount now.

Our job is to determine how much Mr. Cash should get. We reason as follows:

If Mr. Cash accepts P dollars, then the P dollars deposited at 8% for 20 years should yield the same amount as the $1,000 monthly payments for 20 years. In other words, we are comparing the future values for both Mr. Cash and Mr. Credit, and we would like the future values to equal.

Since Mr. Cash is receiving a lump sum of x dollars, its future value is given by the lump sum formula we studied in Section 6.2, and it is

$$\mathrm{A} = \mathrm{P}(1 + .08/12)^{240}$$

Since Mr. Credit is receiving a sequence of payments, or an annuity, of $1,000 per month, its future value is given by the annuity formula we learned in Section 6.3. This value is

$$\mathrm{A} = \frac{\$1000\left[(1 + .08/12)^{240} - 1\right]}{.08/12}$$

The only way Mr. Cash will agree to the amount he receives is if these two future values are equal. So we set them equal and solve for the unknown.

$$\begin{aligned}
&\mathrm{P}(1 + .08/12)^{240} = \frac{\$1000\left[(1 + .08/12)^{240} - 1\right]}{.08/12} \\
&\mathrm{P}(4.9268) = \$1000(589.02041) \\
&\mathrm{P}(4.9268) = \$589020.41 \\
&\mathrm{P} = \$119,554.36
\end{aligned}$$

The present value of an ordinary annuity of $1,000 each month for 20 years at 8% is $119,554.36

The reader should also note that if Mr. Cash takes his lump sum of P = $119,554.36 and invests it at 8% compounded monthly, he will have an accumulated value of A=$589,020.41 in 20 years.

Which equation to use?

When presented with a finance problem (on an exam or in real life), you're usually not told what type of problem it is or which equation to use. Here are some hints on deciding which equation to use based on the wording of the problem.

The easiest types of problem to identify are loans. Loan problems almost always include words like: "loan", "amortize" (the fancy word for loans), "finance (a car)", or "mortgage" (a home loan). Look for these words. If they're there, you're probably looking at a loan problem. To make sure, see if you're given what your monthly (or annual) payment is, or if you're trying to find a monthly payment.

If the problem is not a loan, the next question you want to ask is: "Am I putting money in an account and letting it sit, or am I making regular (monthly/annually/quarterly) payments or withdrawals?" If you're letting the money sit in the account with nothing but interest changing the balance, then you're looking at a compound interest problem. The exception would be bonds and other investments where the interest is not reinvested; in those cases you're looking at simple interest.

If you're making regular payments or withdrawals, the next questions is: "Am I putting money into the account, or am I pulling money out?" If you're putting money into the account on a regular basis (monthly/annually/quarterly) then you're looking at a basic Annuity problem. Basic annuities are when you are saving money. Usually in an annuity problem, your account starts empty, and has money in the future.

If you're pulling money out of the account on a regular basis, then you're looking at a Payout Annuity problem. Payout annuities are used for things like retirement income, where you start with money in your account, pull money out on a regular basis, and your account ends up empty in the future.

Remember, the most important part of answering any kind of question, money or otherwise, is first to correctly identify what the question is really asking, and to determine what approach will best allow you to solve the problem.

? Try it Now 5

For each of the following scenarios, determine if it is a compound interest problem, a savings annuity problem, a payout annuity problem, or a loans problem. Then solve each problem.

a. Marcy received an inheritance of $20,000, and invested it at 6% interest. She is going to use it for college, withdrawing money for tuition and expenses each quarter. How much can she take out each quarter if she has 3 years of school left?
b. Paul wants to buy a new car. Rather than take out a loan, he decides to save $200 a month in an account earning 3% interest compounded monthly. How much will he have saved up after 3 years?
c. Keisha is managing investments for a non-profit company. They want to invest some money in an account earning 5% interest compounded annually with the goal to have $30,000 in the account in 6 years. How much should Keisha deposit into the account?
d. Miao is going to finance new office equipment at a 2% rate over a 4 year term. If she can afford monthly payments of $100, how much new equipment can she buy?
e. How much would you need to save every month in an account earning 4% interest to have $5,000 saved up in two years?

Answer

a. This is a payout annuity problem. She can pull out $1833.60 a quarter.
b. This is a savings annuity problem. He will have saved up $7,524.11.
c. This is compound interest problem. She would need to deposit $22,386.46.
d. This is a loans problem. She can buy $4,609.33 of new equipment.
e. This is a savings annuity problem. You would need to save $200.46 each month

---

CHAPTER OVERVIEW

9: Geometry

Thumbnail: A two-dimensional perspective projection of a sphere (CC BY-3.0; Geek3 via Wikipedia).

9.1: Lines and Angles

You use geometric terms in everyday language, often without thinking about it. For example, any time you say “walk along this line” or “watch out, this road quickly angles to the left” you are using geometric terms to make sense of the environment around you. You use these terms flexibly, and people generally know what you are talking about. In the world of mathematics, each of these geometric terms has a specific definition. It is important to know these definitions—as well as how different figures are constructed—to become familiar with the language of geometry. Let’s start with a basic geometric figure: the plane.

Plane

A **plane** is a flat surface that continues forever (or, in mathematical terms, infinitely) in every direction. It has two dimensions: length and width.

You can visualize a plane by placing a piece of paper on a table. Now imagine that the piece of paper stays perfectly flat and extends as far as you can see in two directions, left-to-right and front-to-back. This gigantic piece of paper gives you a sense of what a geometric plane is like: it continues infinitely in two directions. (Unlike the piece of paper example, though, a geometric plane has no height.) A plane can contain a number of geometric figures.

Point

The most basic geometric idea is a **point**, which has no dimensions. A point is simply a location on the plane. It is represented by a dot. Three points that don’t lie in a straight line will determine a plane.

The image below shows four points, labeled *A*, *B*, *C*, and *D*.

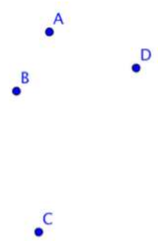


Figure 9.1.1: A set of points

Line

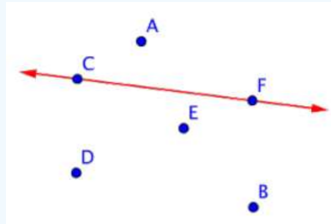
Two points on a plane determine a **line**. A line is a one-dimensional figure that is made up of an infinite number of individual points placed side by side. In geometry, all lines are assumed to be straight; if they bend they are called a curve. A line continues infinitely in two directions.

Below is an image of line *AB* or, in geometric notation, $\overleftrightarrow{AB}$. The arrows indicate that the line keeps going forever in the two directions. This line could also be called line *BA*. While the order of the points does not matter for a line, it is customary to name the two points in alphabetical order.

Figure 9.1.2: Line *AB*

Example 1

Name the line shown in red.



Solution

The red line goes through the points C and F, so the line is $\overleftrightarrow{CF}$.

Line Segment

The section between any two points on a line is called a **line segment**.

A line segment can be very long, very short, or somewhere in between. The difference between a **line** and a **line segment** is that the line segment has two endpoints and a line goes on forever. A line segment is denoted by its two endpoints, as in $\overline{CD}$ in the image below.

Figure 9.1.3: Line segment *CD*

Ray

A **ray** has one endpoint and goes on forever in one direction.

Mathematicians name a ray with notation like $\overrightarrow{EF}$, where point E is the endpoint and F is a point on the ray. When naming a ray, we always say the endpoint first. Note that $\overrightarrow{FE}$ would have the endpoint at F, and continue through E, which is a different ray than $\overrightarrow{EF}$, which would have an endpoint at E, and continue through F.

The term "ray" may be familiar because it is a common word in English. "Ray" is often used when talking about light. While a ray of light resembles the geometric term "ray," it does not go on forever, and it has some width. A geometric ray has no width; only length.

Below is an image of ray *EF* or $\overrightarrow{EF}$. Notice that the endpoint is E.

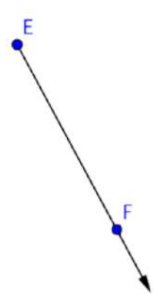
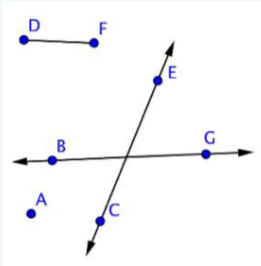


Figure 9.1.4: Ray *EF*

Example 2

Identify each line and line segment in the picture below.



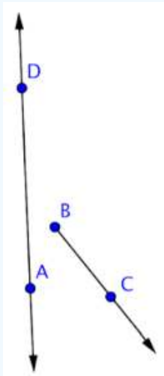
Solution

Two points define a line, and a line is denoted with arrows. There are two lines in this picture: $\overleftrightarrow{CE}$ and $\overleftrightarrow{BG}$.

A line segment is a section between two points. $\overline{DF}$ is a line segment. But there are also two more line segments on the lines themselves: $\overline{CE}$ and $\overline{BG}$.

Example 3

Identify each point and ray in the picture below.



Solution

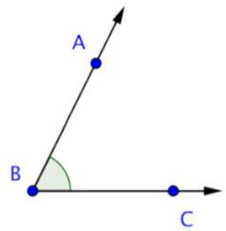
There are four points: A, B, C, and D. There are also three rays, though only one may be obvious.

Ray $\overrightarrow{BC}$ begins at point B and goes through C. Two more rays exist on line $\overleftrightarrow{AD}$: they are $\overrightarrow{DA}$ and $\overrightarrow{AD}$.

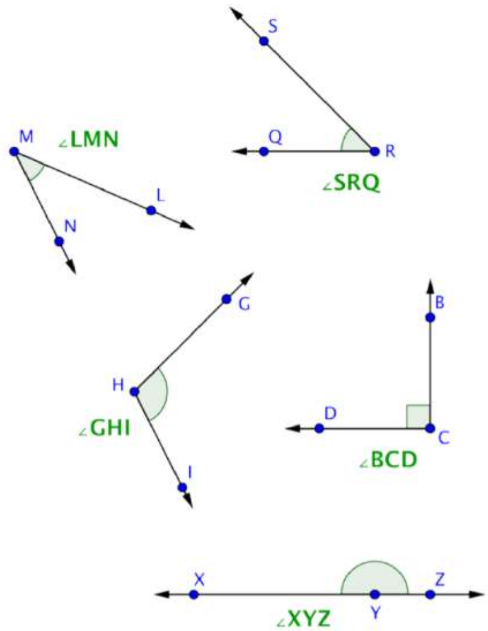
Angle

Two rays with a common endpoint make up an **angle**. The common endpoint of the angle is called the **vertex**.

The angle ABC is shown below. This angle can also be called ∠ABC, ∠CBA or simply ∠B. When you are naming angles, be careful to include the vertex (here, point B) as the middle letter.



The image below shows a few angles on a plane. Notice that the label of each angle is written “point-vertex-point,” and the geometric notation is in the form ∠ABC.



Sometimes angles are very narrow; sometimes they are very wide. When people talk about the “size” of an angle, they are referring to the arc between the two rays. The length of the rays has nothing to do with the size of the angle itself. Drawings of angles will often include an arc (as shown above) to help the reader identify the correct ‘side’ of the angle.

Think about an analog clock face. The minute and hour hands are both fixed at a point in the middle of the clock. As time passes, the hands rotate around the fixed point, making larger and smaller angles as they go. The length of the hands does not impact the angle that is made by the hands.

An angle is measured in degrees, represented by the symbol º. A circle is defined as having 360º. (In skateboarding and basketball, “doing a 360” refers to jumping and doing one complete body rotation.

Types of Angles

A **right angle** is any degree that measures exactly 90º. This represents exactly one-quarter of the way around a circle. Rectangles contain exactly four right angles. A corner mark is often used to denote a right angle, as shown in right angle DCB below.

Angles that are between 0º and 90º (smaller than right angles) are called **acute angles**.

Angles that are between 90º and 180º (larger than right angles and less than 180º) are called **obtuse angles**.

And an angle that measures exactly 180º is called a **straight angle** because it forms a straight line.

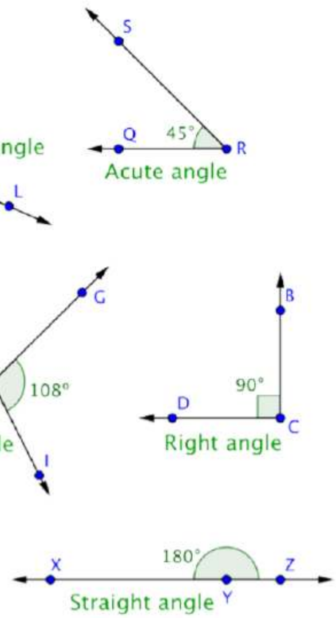
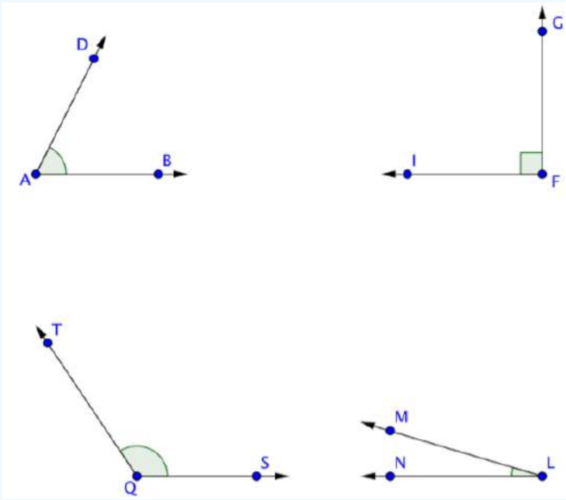


Figure 9.1.5: Examples of Angles

✔ Example 4

Label each angle below as acute, right, or obtuse.



Solution

You can start by identifying any right angles.

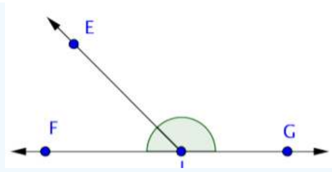
∠GFI is a right angle, as indicated by the corner mark at vertex F.

Acute angles will be smaller than ∠GFI (or less than 90º). This means that ∠DAB and ∠MLN are acute.

∠TQS is larger than ∠GFI, so it is an obtuse angle.

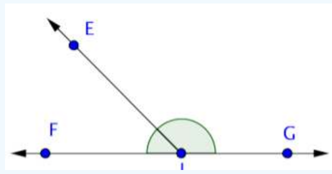
✔ Example 5

Identify each point, ray, and angle in the picture below.

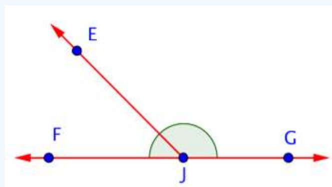


Solution

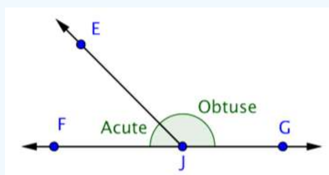
Begin by identifying each point in the figure. There are 4: E, F, G, and J.



Now find rays. A ray begins at one point, and then continues through another point towards infinity (indicated by an arrow). Three rays start at point *J*: $\overrightarrow{JE}$, $\overrightarrow{JF}$, and $\overrightarrow{JG}$. But also notice that a ray could start at point *F* and go through *J* and *G*, and another could start at point *G* and go through *J* and *F*. These rays can be represented by $\overrightarrow{GF}$ and $\overrightarrow{FG}$.



Finally, look for angles. ∠EJG is obtuse, ∠EJF is acute, and ∠FJG is straight. (Don't forget those straight angles!)



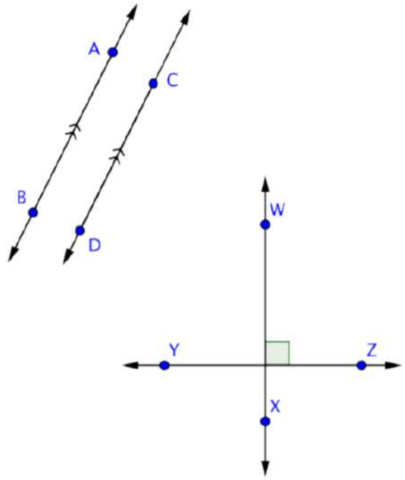
Imagine two separate and distinct lines on a plane. There are two possibilities for these lines: they will either intersect at one point, or they will never intersect. When two lines intersect, four angles are formed. Understanding how these angles relate to each other can help you figure out how to measure them, even if you only have information about the size of one angle.

Parallel and Perpendicular

Parallel lines are two or more lines that never intersect.

Perpendicular lines are two lines that intersect at a 90º (right) angle.

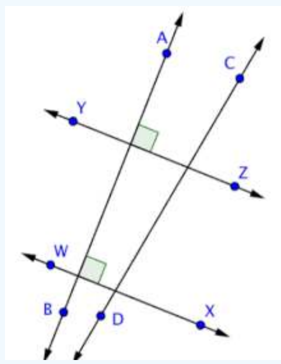
The image below shows some parallel and perpendicular lines. The geometric symbol for parallel is ||, so you can show that *AB* || *CD*. Parallel lines are also often indicated by the marking >> on each line (or just a single > on each line). Perpendicular lines are indicated by the symbol ⊥, so you can write $\overleftrightarrow{WX} \perp \overleftrightarrow{YZ}$.



If two lines are parallel, then any line that is perpendicular to one line will also be perpendicular to the other line. Similarly, if two lines are both perpendicular to the same line, then those two lines are parallel to each other. Let's take a look at one example and identify some of these types of lines.

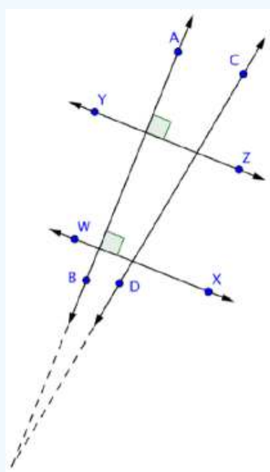
Example 6

Identify a set of parallel lines and a set of perpendicular lines in the image below.

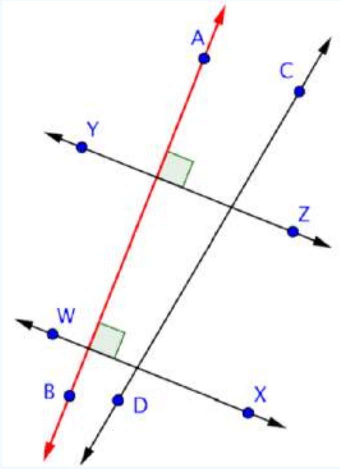


Solution

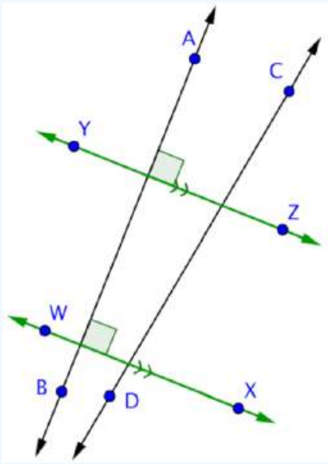
Parallel lines never meet, and perpendicular lines intersect at a right angle. $\overleftrightarrow{AB}$ and $\overleftrightarrow{CD}$ do not intersect in this image, but if you imagine extending both lines, they will intersect soon. So, they are neither parallel nor perpendicular.



$\overleftrightarrow{AB}$ is perpendicular to both $\overleftrightarrow{WX}$ and $\overleftrightarrow{YZ}$, as indicated by the right-angle marks at the intersection of those lines.



Since $\overleftrightarrow{AB}$ is perpendicular to both lines, then $\overleftrightarrow{WX}$ and $\overleftrightarrow{YZ}$ are parallel.



$\overleftrightarrow{WX} \parallel \overleftrightarrow{YZ}$

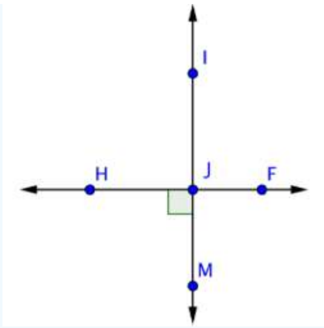
$\overleftrightarrow{AB} \perp \overleftrightarrow{WX}$, $\overleftrightarrow{AB} \perp \overleftrightarrow{YZ}$

Understanding how parallel and perpendicular lines relate can help you figure out the measurements of some unknown angles. To start, all you need to remember is that perpendicular lines intersect at a 90º angle and that a straight angle measures 180º.

The measure of an angle such as ∠A is written as m∠A. Look at the example below. How can you find the measurements of the unmarked angles?

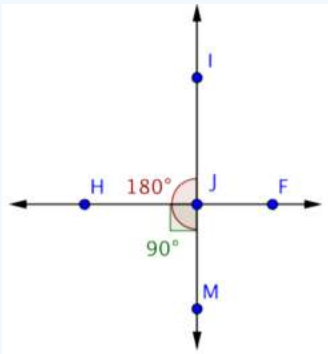
Example 7

Find the measurement of ∠IJF.



Solution

Only one angle, ∠HJM, is marked in the image. Notice that it is a right angle, so it measures 90º. ∠HJM is formed by the intersection of lines $\overleftrightarrow{IM}$ and $\overleftrightarrow{HF}$. Since $\overleftrightarrow{IM}$ is a line, ∠IJM is a straight angle measuring 180º.

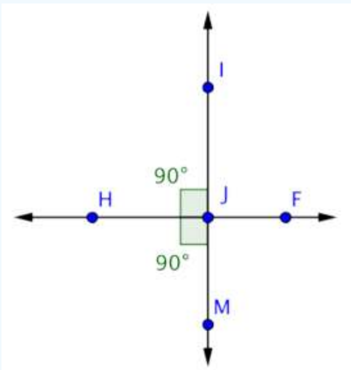


You can use this information to find the measurement of ∠HJI :

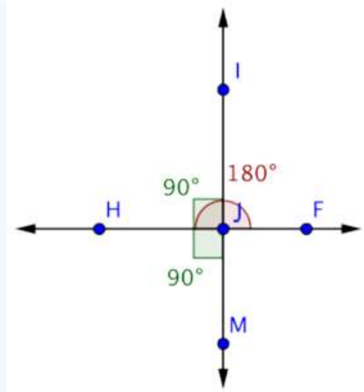
m∠HJM + m∠HJI = m∠IJM

90º + m∠HJI = 180º

m∠HJI = 90º



Now use the same logic to find the measurement of ∠IJF. ∠IJF is formed by the intersection of lines $\overleftrightarrow{IM}$ and $\overleftrightarrow{HF}$. Since $\overleftrightarrow{HF}$ is a line, ∠HJF will be a straight angle measuring 180º.

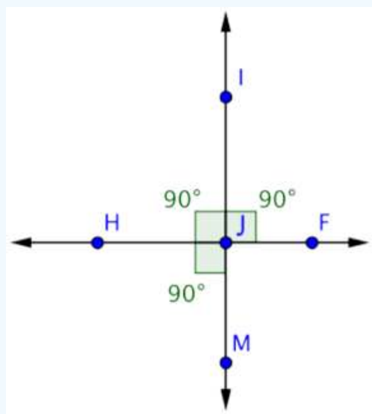


You know that ∠HJI measures 90º. Use this information to find the measurement of ∠IJF:

m∠HJM + m∠IJF = m∠HJF

90º + m∠IJF = 180º

m∠IJF = 90º



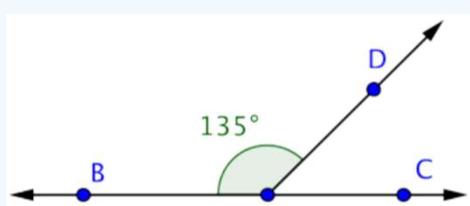
Thus, m∠IJF = 90º

In this example, you may have noticed that angles ∠HJI, ∠IJF, and ∠HJM are all right angles. (If you were asked to find the measurement of ∠FJM, you would find that angle to be 90º, too.) This is what happens when two lines are perpendicular—the four angles created by the intersection are all right angles.

Not all intersections happen at right angles, though. In the example below, notice how you can use the same technique as shown above (using straight angles) to find the measurement of a missing angle.

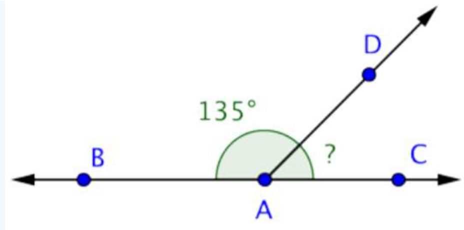
✓ Example 8

Find the measurement of ∠DAC.

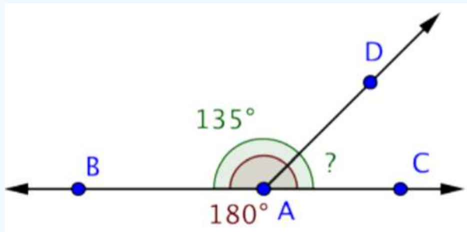


Solution

This image shows the line $\overleftrightarrow{BC}$ and the ray $\overrightarrow{AD}$ intersecting at point A. The measurement of ∠BAD is 135º. You can use straight angles to find the measurement of ∠DAC.



∠BAC is a straight angle, so it measures 180º.

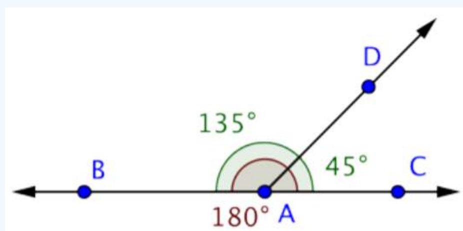


Use this information to find the measurement of ∠DAC.

m∠BAD + m∠DAC = m∠BAC

135º + m∠DAC = 180º

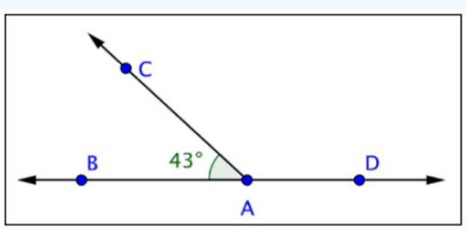
m∠DAC = 45º



Thus, m∠DAC = 45º

Try it Now 1

Find the measurement of ∠CAD.



Answer

m∠CAD = 137º

Supplementary and Complementary

Two angles whose measures add up to 180º are called **supplementary angles**.

Two angles whose measurements add up to 90º, they are called **complementary angles**.

Hint: One way to remember the difference between the two terms is that "corner" and "complementary" each begin with *c* (a 90º angle looks like a corner), while straight and "supplementary" each begin with *s* (a straight angle measures 180º).

If you can identify supplementary or complementary angles within a problem, finding missing angle measurements is often simply a matter of adding or subtracting.

Example 9

Two angles are supplementary. If one of the angles measures 48º, what is the measurement of the other angle?

Solution

Two supplementary angles make up a straight angle, so the measurements of the two angles will be 180º.

$$m\angle A + m\angle B = 180^\circ$$

You know the measurement of one angle. To find the measurement of the second angle, subtract 48º from 180º.

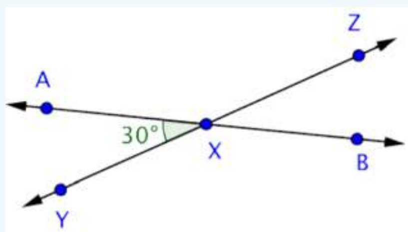
$$48^\circ + m\angle B = 180^\circ$$

$$m\angle B = 180^\circ - 48^\circ$$

$$m\angle B = 132^\circ$$

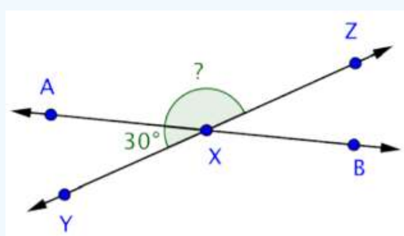
Example 10

Find the measurement of ∠AXZ.



Solution

This image shows two intersecting lines, $\overleftrightarrow{AB}$ and $\overleftrightarrow{YZ}$. They intersect at point X, forming four angles. Angles ∠AXY and ∠AXZ are supplementary because together they make up the straight angle ∠YXZ.

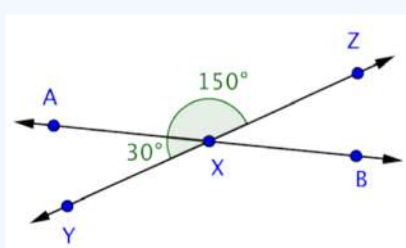


Use this information to find the measurement of ∠AXZ.

$$m\angle AXY + m\angle AXZ = m\angle YXZ$$

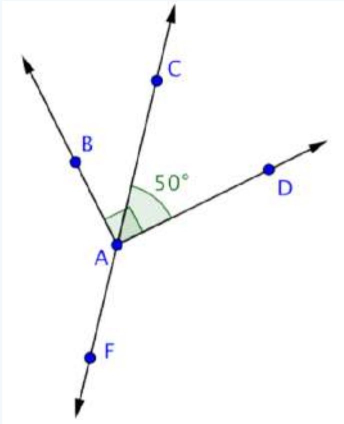
$$30^\circ + m\angle AXZ = 180^\circ$$

$$m\angle AXZ = 150^\circ$$



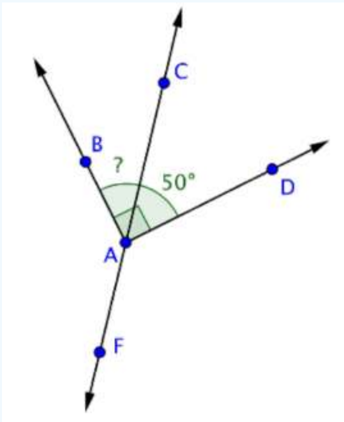
Example 11

Find the measurement of ∠BAC.



Solution

This image shows the line $\overleftrightarrow{CF}$ and the rays $\overleftrightarrow{AB}$ and $\overleftrightarrow{AD}$, all intersecting at point A. Angle ∠BAD is a right angle. Angles ∠BAC and ∠CAD are complementary because together they create ∠BAD.

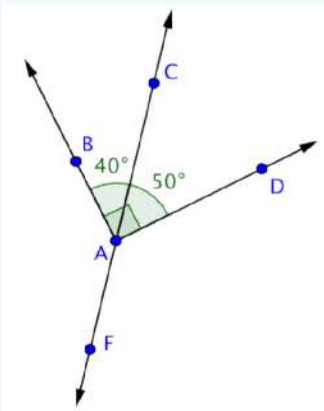


Use this information to find the measurement of ∠BAC .

$$m\angle BAC + m\angle CAD = m\angle BAD$$

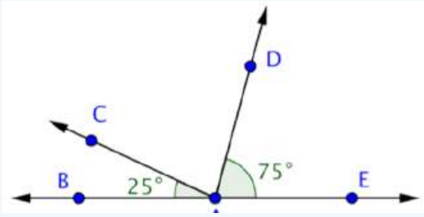
$$m\angle BAC + 50^\circ = 90^\circ$$

$$m\angle BAC = 40^\circ$$



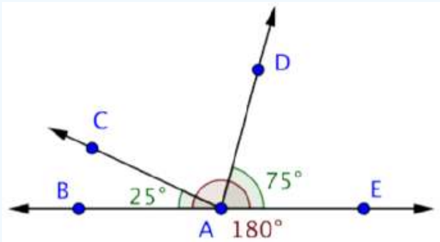
Example 12

Find the measurement of ∠CAD.



Solution

You know the measurements of two angles here: ∠CAB and ∠DAE. You also know that m∠BAE = 180º.



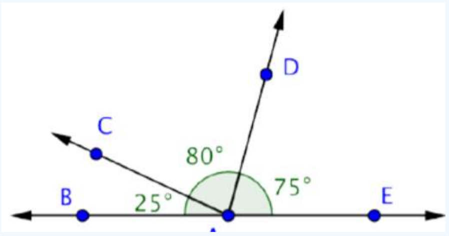
Use this information to find the measurement of ∠CAD.

$$m\angle BAC + m\angle CAD + m\angle DAE = m\angle BAE$$

$$25^{o} + m\angle CAD + 75^{o} = 180^{o}$$

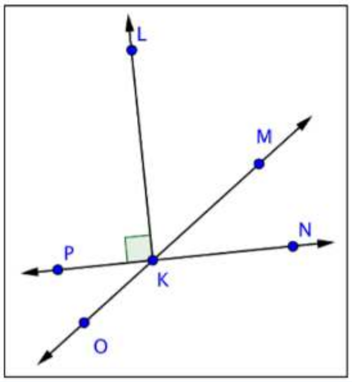
$$m\angle CAD + 100^{o} = 180^{o}$$

$$m\angle CAD = 80^{o}$$



Try it Now 2

Which pair of angles is complementary?



A) ∠PKO and ∠MKN

B) ∠PKO and ∠PKM

C) ∠LKP and ∠LKN

D) ∠LKM and ∠MKN

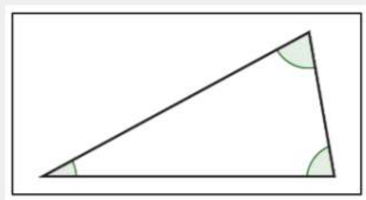
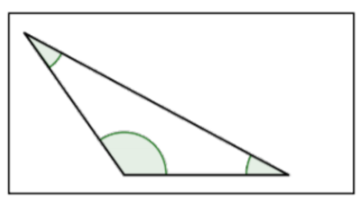
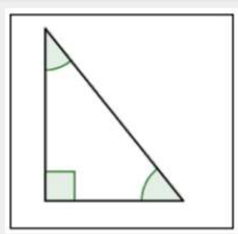
Answer

D) ∠LKM and ∠MKN

9.2: Polygons

> ✎ **Polygon**
>
> A **polygon** is a closed plane figure with three or more straight sides called **edges**. An endpoint of an edge a **vertex**. A polygon is **regular** if all of its edges are the same length and all of its angles have the same measure.
>
> Polygons each have a special name based on the number of sides they have. For example, the polygon with three sides is called a triangle because “tri” is a prefix that means “three.” Its name also indicates that this polygon has three angles. The prefix “poly” means many.

The smallest polygon you can have is a triangle, which has three sides. The table below shows and describes three classifications of triangles. Notice how the types of angles in the triangle are used to classify the triangle.

Name of Triangle	Picture of Triangle	Description
Acute Triangle		A triangle with 3 acute angles (3 angles measuring between 0° and 90°).
Obtuse Triangle		A triangle with 1 obtuse angle (1 angle measuring between 90° and 180°).
Right Triangle		A triangle containing one right angle (1 angle that measures 90°). Note that the right angle is shown with a corner mark and does not need to be labeled 90°.

> ✎ **Sum of the Angles in a Triangle**
>
> The sum of the measures of the three interior angles of a triangle is always 180°.
>
> This fact can be applied to find the measure of the third angle of a triangle, if you are given the other two.

> ✔ **Example 1**
>
> A triangle has two angles that measure 35° and 75°. Find the measure of the third angle.
>
> **Solution**
>
> The sum of the three interior angles of a triangle is 180°.
>
> $$35° + 75° + x = 180°$$
>
> Find the value of x.
>
> $$110° + x = 180°$$
>
> $$x = 180° - 110°$$

$$x = 70°$$

Answer: The third angle of the triangle measures 70°.

Example 2

One of the angles in a right triangle measures 57º. Find the measurement of the third angle.

Solution

The sum of the three angles of a triangle is 180°. One of the angles has a measure of 90° as it is a right triangle.

$$57° + 90° + x = 180°$$

Simplify.

$$147º + x = 180°$$

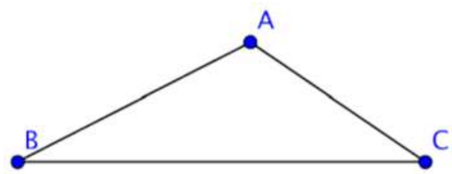
Find the value of x.

$$x = 180º - 147º$$

$$x = 33º$$

Answer: The third angle of the right triangle measures 33°.

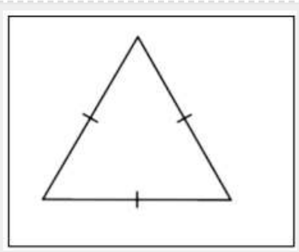
There is an established convention for naming triangles. The labels of the vertices of the triangle, which are generally capital letters, are used to name a triangle.

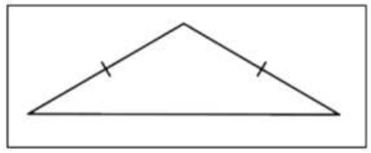
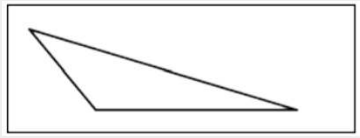


You can call this triangle *ABC* or ΔABC since *A*, *B*, and *C* are vertices of the triangle. When naming the triangle, you can begin with any vertex. Then keep the letters in order as you go around the polygon. The triangle above could be named in a variety of ways: ΔABC, or ΔCBA. The sides of the triangle are line segments *AB*, *AC*, and *CB*.

Just as triangles can be classified as acute, obtuse, or right based on their angles, they can also be classified by the length of their sides. Sides of equal length are called **congruent** sides. While we designate a segment joining points *A* and *B* by the notation $\overline{AB}$, we designate the length of a segment joining points *A* and *B* by the notation *AB* without a segment bar over it. The length *AB* is a number, and the segment $\overline{AB}$ is the collection of points that make up the segment.

Mathematicians show congruency by putting a hash mark symbol through the middle of sides of equal length. If the hash mark is the same on one or more sides, then those sides are congruent. If the sides have different hash marks, they are *not* congruent. The table below shows the classification of triangles by their side lengths.

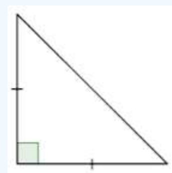
Name of Triangle	Picture of Triangle	Description
Equilateral Triangle		A triangle whose three sides have the same length. These sides of equal length are called congruent sides.

Isosceles Triangle		A triangle with exactly two congruent sides.
Scalene Triangle		A triangle in which all three sides are a different length.

To describe a triangle even more specifically, you can use information about both its sides and its angles.

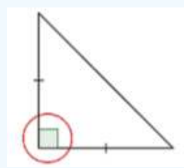
Example 3

Classify the triangle below.

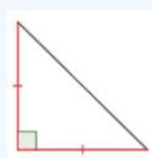


Solution

Notice what kind of angles the triangle has. Since one angle is a right angle, this is a right triangle.



Notice the lengths of the sides. Are there congruence marks or other labels?

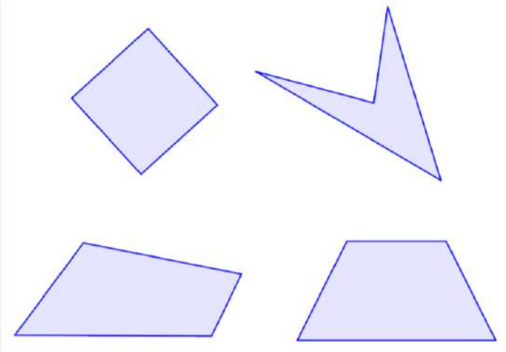


The congruence marks tell us there are two sides of equal length. So, this is an isosceles triangle.

Answer: This is an isosceles right triangle

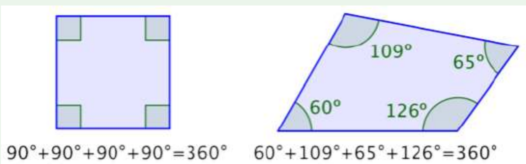
Quadrilaterals

Quadrilaterals are polygons with exactly four side.



Sum of the Angles in a Quadrilateral

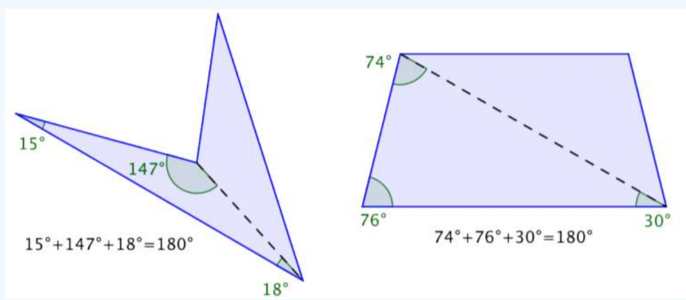
The sum of the interior angles of any quadrilateral is 360°.



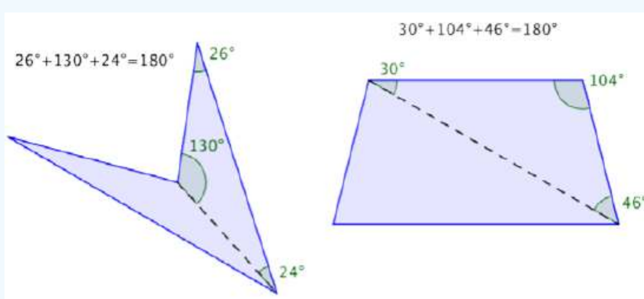
Example 4

You can also use your knowledge of triangles as a way to understand why the sum of the interior angles of any quadrilateral is 360°. Any quadrilateral can be divided into two triangles as shown in the images below.

In the first image, the quadrilaterals have each been divided into two triangles. The angle measurements of one triangle are shown for each.



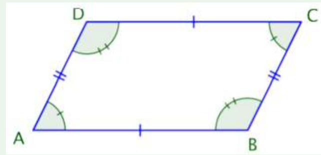
These measurements add up to 180º. Now look at the measurements for the other triangles—they also add up to 180º!



Since the sum of the interior angles of any triangle is 180° and there are two triangles in a quadrilateral, the sum of the angles for each quadrilateral is 360°.

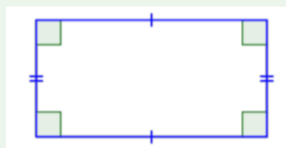
Types of Quadrilaterals

Let’s start by examining the group of quadrilaterals that have two pairs of parallel sides. These quadrilaterals are called **parallelograms**. They take a variety of shapes, but one classic example is shown below.

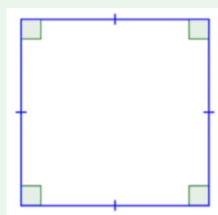


Imagine extending the pairs of opposite sides. They would never intersect because they are parallel. Notice, also, that the opposite angles of a parallelogram are congruent, as are the opposite sides. (Remember that “congruent” means “the same size.”) The geometric symbol for congruent is $\cong$, so you can write ∠A ≅ ∠C and ∠B ≅ ∠D. The parallel sides are also the same length: $\overline{AB} \cong \overline{DC}$ and $\overline{BC} \cong \overline{AD}$. These relationships are true for all parallelograms.

There are two special cases of parallelograms that will be familiar to you from your earliest experiences with geometric shapes. The first special case is called a **rectangle**. By definition, a rectangle is a parallelogram because its pairs of opposite sides are parallel. A rectangle also has the special characteristic that all of its angles are right angles; all four of its angles are congruent.

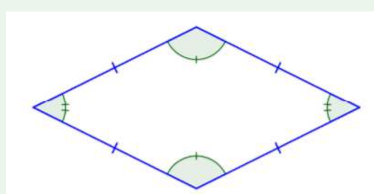


The other special case of a parallelogram is a special type of rectangle, a **square**. A square is one of the most basic geometric shapes. It is a special case of a parallelogram that has four congruent sides and four right angles.



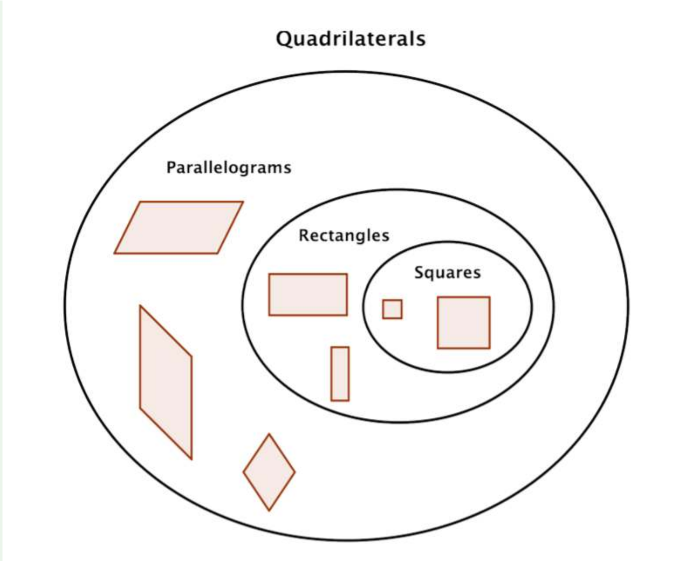
A square is also a rectangle because it has two sets of parallel sides and four right angles. A square is also a parallelogram because its opposite sides are parallel. So, a square can be classified in any of these three ways, with “parallelogram” being the least specific description and “square,” the most descriptive.

Another quadrilateral that you might see is called a **rhombus**. All four sides of a rhombus are congruent. Its properties include that each pair of opposite sides is parallel, also making it a parallelogram.

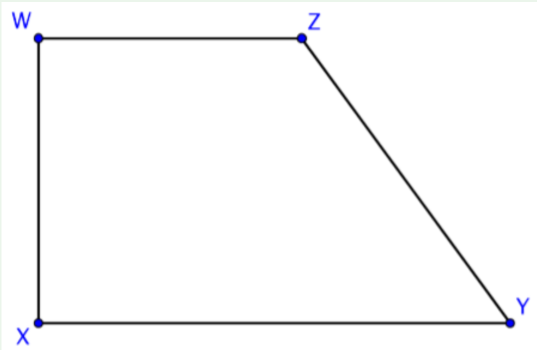


In summary, all squares are rectangles, but not all rectangles are squares. All rectangles are parallelograms, but not all parallelograms are rectangles. And all of these shapes are quadrilaterals.

The diagram below illustrates the relationship between the different types of quadrilaterals.

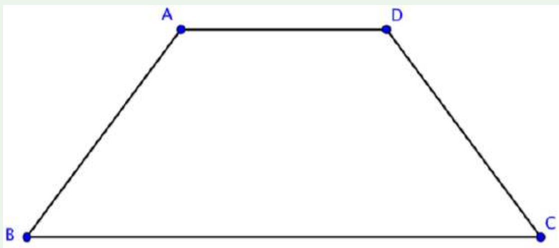


There is another special type of quadrilateral. This quadrilateral has the property of having only one pair of opposite sides that are parallel. Here is one example of a **trapezoid**.



Notice that $\overline{XY} || \overline{WZ}$, and that $\overline{WX}$ and $\overline{ZY}$ are not parallel. You can easily imagine that if you extended sides $\overline{WX}$ and $\overline{ZY}$, they would intersect above the figure.

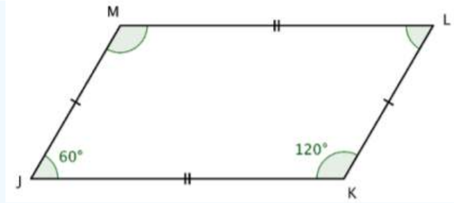
If the non-parallel sides of a trapezoid are congruent, the trapezoid is called an **isosceles trapezoid**. Like the similarly named triangle that has two sides of equal length, the isosceles trapezoid has a pair of opposite sides of equal length. The other pair of opposite sides is parallel. Below is an example of an isosceles trapezoid.



In this trapezoid *ABCD*, $\overline{BC} || \overline{AD}$ and $\overline{AB} \cong \overline{CD}$.

Example 5

Determine the measures of ∠M and ∠L.



Solution

Identify opposite angles.

$$\angle L \text{ is opposite } \angle J$$

$$\angle M \text{ is opposite } \angle K$$

A property of parallelograms is that opposite angles are congruent.

$$\angle L \cong \angle J$$

$$\angle M \cong \angle K$$

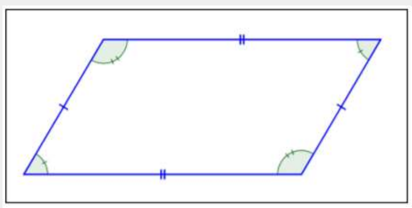
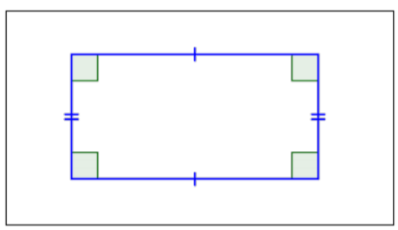
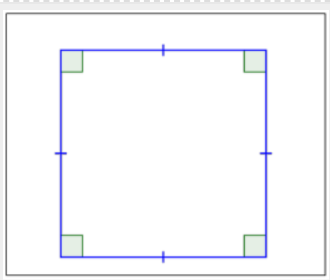
Use the given angle measurements to determine measures of opposite angles.

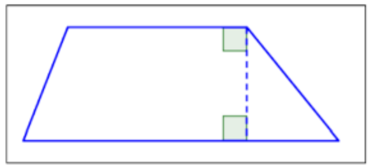
$$m\angle J = 60°, \text{ so } m\angle L = 60°$$

$$m\angle K = 120°, \text{ so } m\angle M = 120°$$

Answer: $m\angle L = 60°$ and $m\angle M = 120°$

The table below summarizes the special types of quadrilaterals and some of their properties.

Name of Quadrilateral	Quadrilateral	Description
Parallelogram		2 pairs of parallel sides. Opposite sides and opposite angles are congruent.
Rectangle		2 pairs of parallel sides. 4 right angles (90°). Opposite sides are parallel and congruent. All angles are congruent.
Square		4 congruent sides. 4 right angles (90°). Opposite sides are parallel. All angles are congruent.

Trapezoid		Only one pair of opposite sides is parallel.

Sum of the Angles in a Polygon

The sum of the measures of the interior angles of a polygon having *n* sides is $(n-2)\times 180^{\circ}$.

If the polygon is a regular polygon, then each interior angle is $\frac{(n-2)\times 180^{\circ}}{n}$

Example 6

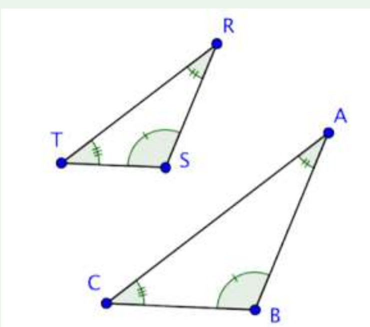
What is the measure of an interior angle of a regular 12-sided polygon (dodecagon)?

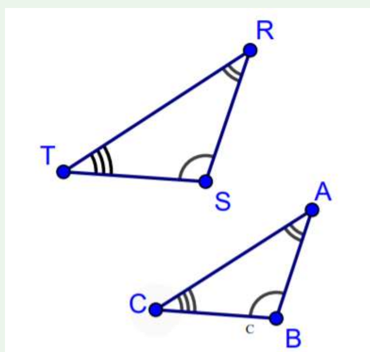
Solution

Since the polygon is regular, each interior angle will have the same measure of $\frac{(12-2)\times 180^{\circ}}{12}$ = 150°

Similar Polygon

Two polygons are **similar** if their corresponding sides are proportional and their corresponding angles are equal.



OR


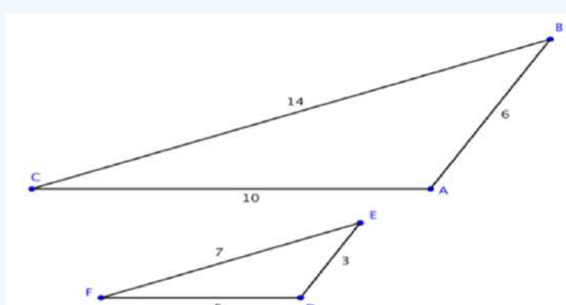
$\frac{\overline{AB}}{\overline{RS}}=\frac{\overline{BC}}{\overline{ST}}=\frac{\overline{AC}}{\overline{RT}}$

∠A ≅ ∠R, ∠B ≅ ∠S, and ∠C ≅ ∠T

This extends to all polygons and not *just* triangles.

Example 7

Determine if the triangles below are similar by seeing if their corresponding sides are proportional.



Solution

First determine the corresponding sides, which are opposite corresponding angles.

$$\overline{CA} \leftrightarrow \overline{FD}$$

$$\overline{AB} \leftrightarrow \overline{DE}$$

$$\overline{BC} \leftrightarrow \overline{EF}$$

Write the corresponding side lengths as ratios.

$$\frac{\overline{CA}}{\overline{FD}} = \frac{\overline{AB}}{\overline{DE}} = \frac{\overline{BC}}{\overline{EF}}$$

Substitute the side lengths into the ratios, and determine if the ratios of the corresponding sides are equivalent. They are, so the triangles are similar.

$$\frac{10}{5} = \frac{6}{3} = \frac{14}{7}$$

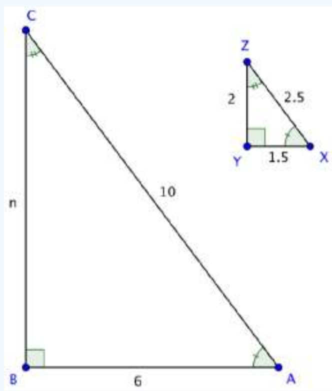
$$2 = 2 = 2$$

Answer: ΔABC and ΔDEF are similar.

The mathematical symbol ~ means "is similar to". So, you can write ΔABC is similar to ΔDEF as ΔABC ~ ΔDEF.

Example 8

ΔABC and ΔXYZ are similar triangles. What is the length of side *BC?*



Solution

In similar triangles, the ratios of corresponding sides are proportional. Set up a proportion of two ratios, one that includes the missing side.

$$\frac{BC}{YZ} = \frac{AB}{XY}$$

Substitute in the known side lengths for the side names in the ratio. Let the unknown side length be *n*.

$$\frac{n}{2} = \frac{6}{1.5}$$

Solve for *n* using cross multiplication.

$$2 \cdot 6 = 1.5 \cdot n$$

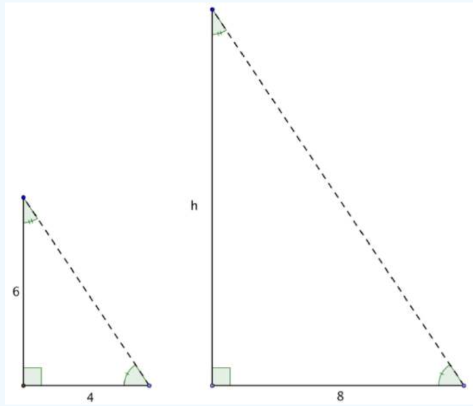
$$12 = 1.5n$$

$$8 = n$$

Applying knowledge of triangles, similarity, and congruence can be very useful for solving problems in real life. Just as you can solve for missing lengths of a triangle drawn on a page, you can use triangles to find unknown distances between locations or objects.

Example 9

When the sun is at a certain angle in the sky, a 6-foot tree will cast a 4-foot shadow. How tall is a tree that casts an 8-foot shadow?



Solution

The angle measurements are the same, so the triangles are similar triangles. Since they are similar triangles, you can use proportions to find the size of the missing side.

$$\frac{\text{Tree 1}}{\text{Tree 2}} = \frac{\text{Shadow 1}}{\text{Shadow 2}}$$

Set up a proportion comparing the heights of the trees and the lengths of their shadows.

Substitute in the known lengths. Call the missing tree height h.

$$\frac{6}{h} = \frac{4}{8}$$

Solve for h using cross-multiplication.

$$6 \cdot 8 = 4h$$

$$48 = 4h$$

$$12 = h$$

Answer: The tree is 12 feet tall.

9.3: Perimeter and Area

Perimeter and area are two important and fundamental mathematical topics. They help you to quantify physical space and also provide a foundation for more advanced mathematics found in algebra, trigonometry, and calculus. Perimeter is a measurement of the distance around a shape and area gives us an idea of how much surface the shape covers.

Knowledge of area and perimeter is applied practically by people on a daily basis, such as architects, engineers, and graphic designers, and is math that is very much needed by people in general. Understanding how much space you have and learning how to fit shapes together exactly will help you when you paint a room, buy a home, remodel a kitchen, or build a deck.

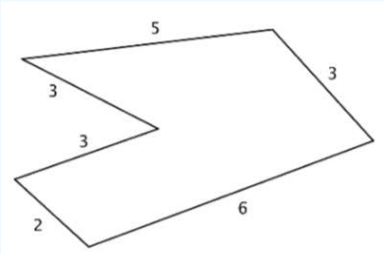
Perimeter and Area

The **perimeter** of a polygon is the sum of the lengths of tthe sides of the polygon.

The **area** is a measure of the amount of surface that polygon covers.

Example 1

Find the perimeter of the given figure. All measurements indicated are inches.



Solution

Since all the sides are measured in inches, just add the lengths of all six sides to get the perimeter.

$$P = 5 + 3 + 6 + 2 + 3 + 3$$

Remember to include units.

Answer: $P = 22$ inches

Example 2

Find the perimeter of a triangle with sides measuring 6 cm, 8 cm, and 12 cm.

Solution

Since all the sides are measured in centimeters, just add the lengths of all three sides to get the perimeter.

$$P = 6 + 8 + 12$$

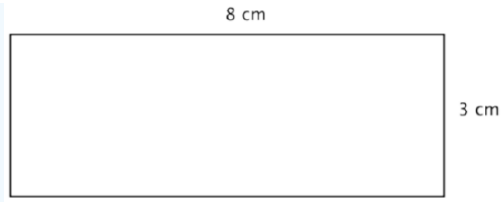
Answer: $P = 26$ centimeters

Perimeter of a Rectangle

If a rectangle has length l and width w, then the perimeter of the rectangle is $P = 2l + 2w$.

Example 3

A rectangle has a length of 8 centimeters and a width of 3 centimeters. Find the perimeter.



Solution

Since this is a rectangle, the opposite sides have the same lengths, 3 cm. and 8 cm. Add up the lengths of all four sides to find the perimeter.

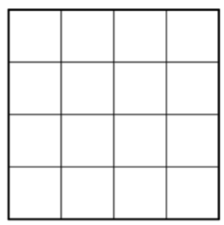
$$P = 2 \times 3 + 2 \times 8$$

Answer $P = 22 \text{ cm}$

Area of Parallelograms

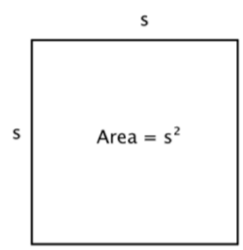
The area of a two-dimensional figure describes the amount of surface the shape covers. You measure area in square units of a fixed size. Examples of square units of measure are square inches, square centimeters, or square miles. When finding the area of a polygon, you count how many squares of a certain size will cover the region inside the polygon.

Let's look at a 4 x 4 square.



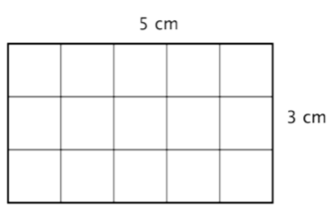
You can count that there are 16 squares, so the area is 16 square units. Counting out 16 squares doesn't take too long, but what about finding the area if this is a larger square or the units are smaller? It could take a long time to count.

Fortunately, you can use multiplication. Since there are 4 rows of 4 squares, you can multiply $4 \cdot 4$ to get 16 squares! And this can be generalized to a formula for finding the area of a square with any length, s: $\text{Area } = s \cdot s = s^2$.



You can write "in^2" for square inches and "ft^2" for square feet.

To help you find the area of the many different categories of polygons, mathematicians have developed formulas. These formulas help you find the measurement more quickly than by simply counting. The formulas you are going to look at are all developed from the understanding that you are counting the number of square units *inside* the polygon. Let's look at a rectangle.



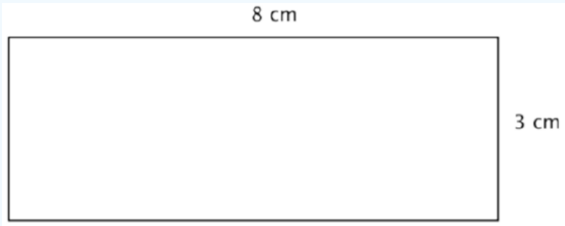
You can count the squares individually, but it is much easier to multiply 3 times 5 to find the number more quickly.

Area of a Rectangle

If a rectangle has length l and width w, then the area of the rectangle is $A = l \times w$.

Example 4

A rectangle has a length of 8 centimeters and a width of 3 centimeters. Find the area.



Solution

Start with the formula for the area of a rectangle, which multiplies the length times the width.

$$A = l \cdot w$$

Substitute 8 for the length and 3 for the width.

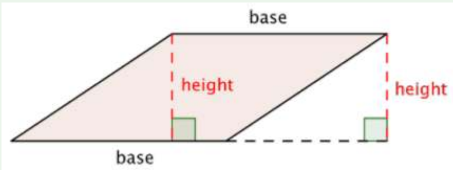
$$A = 8 \cdot 3$$

Be sure to include the units, in this case, square cm.

Answer: $A = 24 \text{ cm}^2$

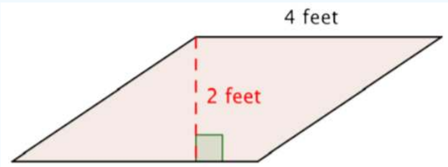
Area of a Parallelogram

The formula for the area of any parallelogram (remember, a rectangle is a type of parallelogram) is the same as that of a rectangle: $\text{Area } = l \cdot w$. Notice in a rectangle, the length and the width are perpendicular. This should also be true for all parallelograms. *Base (b)* for the length (of the base), and *height (h)* for the width of the line perpendicular to the base is often used. So, the formula for a parallelogram is generally written, $\text{Area } = b \cdot h$.



Example 5

Find the area of the parallelogram.



Solution

Start with the formula for the area of a parallelogram: $\text{Area } = \text{base} \cdot \text{height}$.

$$A = b \cdot h$$

Substitute the values into the formula.

$$A = 4 \cdot 2$$

Multiply.

$$A = 8$$

Answer: The area of the parallelogram is 8 ft^2.

Try it Now 1

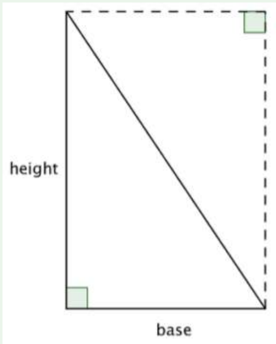
Find the area of a parallelogram with a height of 12 feet and a base of 9 feet.

Answer

The area of the parallelogram is $12 \cdot 9 = 108\ \text{ft}^2$.

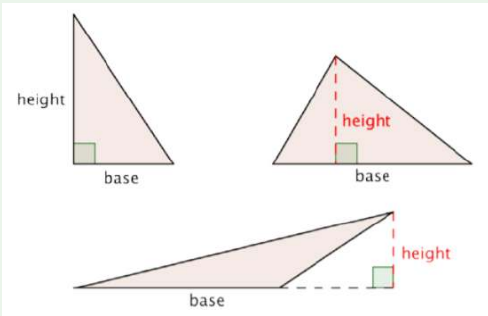
Area of a Triangle

The formula for the area of a triangle can be explained by looking at a right triangle. Look at the image below—a rectangle with the same height and base as the original triangle. The area of the triangle is one half of the rectangle!



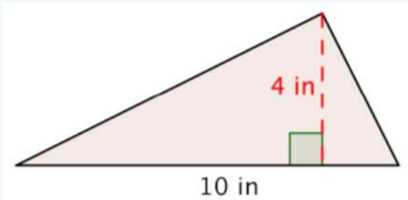
Since the area of two congruent triangles is the same as the area of a rectangle, you can come up with the formula $\text{Area } = \frac{1}{2}bh$ to find the area of a triangle.

When you use the formula for a triangle to find its area, it is important to identify a base and its corresponding height, which is perpendicular to the base.



Example 6

A triangle has a height of 4 inches and a base of 10 inches. Find the area.



Solution

Start with the formula for the area of a triangle.

$$A=\frac{1}{2}bh$$

Substitute 10 for the base and 4 for the height.

$$A=\frac{1}{2}\cdot(10)\cdot(4)$$

Multiply.

$$A=\frac{1}{2}\cdot 40$$

$$A=20$$

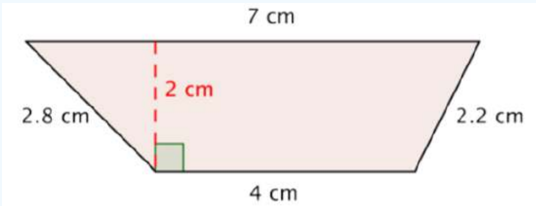
Answer: $A=20\text{ in}^2$

Area of a Trapezoid

To find the area of a trapezoid, take the average length of the two parallel bases and multiply that length by the height: $A=\frac{(b_1+b_2)}{2}h$.

Example 7

Find the area of the trapezoid.



Solution

Start with the formula for the area of a trapezoid.

$$A=\frac{(b_1+b_2)}{2}h$$

Substitute 4 and 7 for the bases and 2 for the height, and find A.

$$A=\frac{(4+7)}{2}\cdot 2$$

$$A=\frac{11}{2}\cdot 2$$

$$A=11$$

Answer: The area of the trapezoid is 11 cm^2

Area Formulas

Use the following formulas to find the areas of different shapes.

square: $A=s^2$

rectangle: $A=l\cdot w$

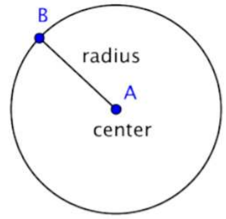
parallelogram: $A=b\cdot h$

triangle: $A=\frac{1}{2}b\cdot h$

trapezoid: $A=\frac{(b_1+b_2)}{2}h$

Circles are a common shape. You see them all over—wheels on a car, Frisbees passing through the air, compact discs delivering data. These are all circles.

A circle represents a set of points, all of which are the same distance away from a fixed, middle point. This fixed point is called the center. The distance from the center of the circle to any point on the circle is called the **radius**. The **diameter** is twice the radius.



The distance around a circle is called the circumference. (Recall, the distance around a polygon is the perimeter.)

One interesting property about circles is that the ratio of a circle's circumference and its diameter is the same for all circles. No matter the size of the circle, the ratio of the circumference and diameter will be the same. The ratio $\frac{C}{d}$ is **pi** and is represented by the Greek letter π.

π is a non-terminating, non-repeating decimal, so it is impossible to write it out completely? The first 10 digits of π are 3.141592653; we will round to 3.14 to approximate pi.

Circumference of a Circle

The circumference C of a circle with radius r is $C = 2\pi r$

Example 8

Find the circumference of a circle with a radius of 2.5 yards.

Solution

To calculate the circumference of a circle given a radius of 2.5 yards, use the formula $C = 2\pi r$. Use 3.14 as an approximation for π.

$$C = 2\pi r$$
$$C = 2\pi \cdot 2.5$$
$$C = \pi \cdot 5$$
$$C \approx 3.14 \cdot 5$$
$$C \approx 15.7$$

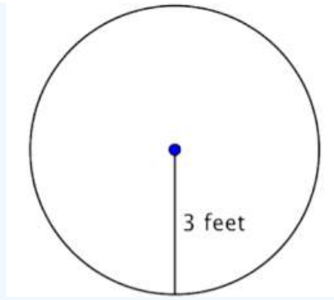
Answer: The circumference is 5π or approximately 15.7 yards.

Area of a Circle

The area A of a circle with radius r is $A = \pi r^2$

Example 9

Find the area of the circle.



Solution

To find the area of this circle, use the formula $A = \pi r^2$. Remember to write the answer in terms of square units, since you are finding the area.

$$A = \pi r^2$$

$$A = \pi \cdot 3^2$$

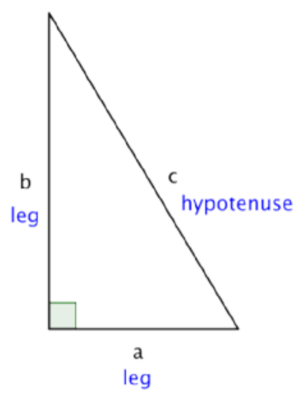
$$A = \pi \cdot 9$$

$$A = 3.14 \cdot 9$$

$$A = 28.26$$

Answer The area is 9π or approximately 28.26 feet^2.

A long time ago, a Greek mathematician named **Pythagoras** discovered an interesting property about **right triangles**: the sum of the squares of the lengths of each of the triangle's **legs** is the same as the square of the length of the triangle's **hypotenuse**. This property—which has many applications in science, art, engineering, and architecture—is now called the **Pythagorean Theorem**.



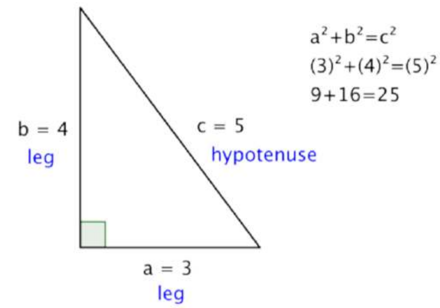
The Pythagorean Theorem

If *a* and *b* are the lengths of the legs of a right triangle and c is the length of the hypotenuse, then the sum of the squares of the lengths of the legs is equal to the square of the length of the hypotenuse.

This relationship is represented by the formula: $a^2 + b^2 = c^2$

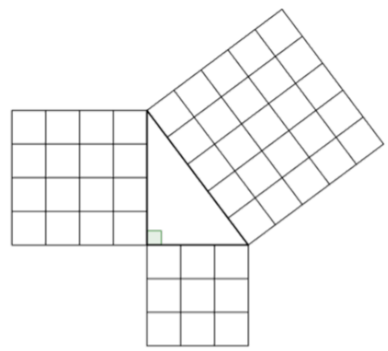
When you see the equation $a^2 + b^2 = c^2$, you can think of this as "the length of side *a* times itself, plus the length of side *b* times itself is the same as the length of side *c* times itself."

Let's try out all of the Pythagorean Theorem with an actual right triangle.



This theorem holds true for this right triangle—the sum of the squares of the lengths of both legs is the same as the square of the length of the hypotenuse. And, in fact, it holds true for all right triangles.

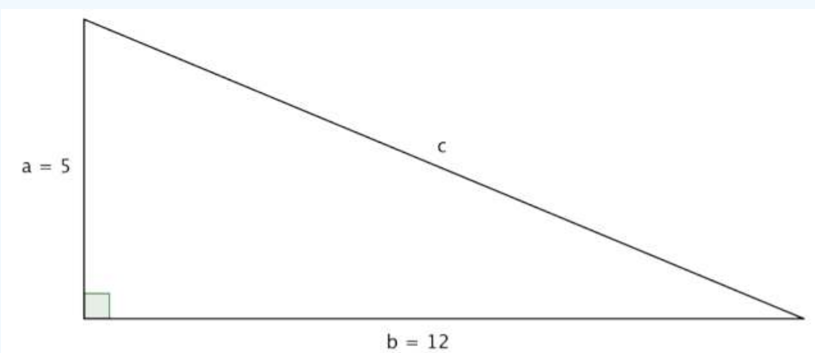
The Pythagorean Theorem can also be represented in terms of area. In any right triangle, the area of the square drawn from the hypotenuse is equal to the sum of the areas of the squares that are drawn from the two legs. You can see this illustrated below in the same 3-4-5 right triangle.



Note that the Pythagorean Theorem only works with *right* triangles.

Example 10

Find the missing length c in this right triangle.



Solution

In the triangle above, you are given measures for legs a and b: 5 and 12, respectively. You can use the Pythagorean Theorem to find a value for the length of c, the hypotenuse.

The Pythagorean Theorem.

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

Substitute known values for a and b.

$$(5)^2+(12)^2=c^2$$

Evaluate.

$$25+144=c^2$$

Simplify. To find the value of c, think about a number that, when multiplied by itself, equals 169. Does 10 work? How about 11? 12? 13? (You can use a calculator to multiply if the numbers are unfamiliar.)

$$169=c^2$$

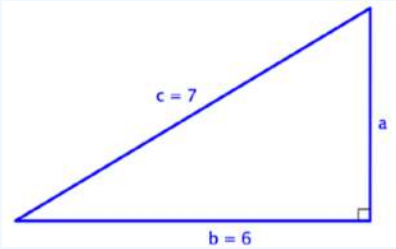
The square root of 169 is 13.

$$c=13$$

Using the formula, you find that the length of c, the hypotenuse, is 13.

Example 11

Find the length of side a in the triangle below. Use a calculator to estimate the square root to one decimal place.



Solution

In this right triangle, you are given the measurements for the hypotenuse, c, and one leg, b. The hypotenuse is always opposite the right angle and it is always the longest side of the triangle.

$$a = ?$$

$$b = 6$$

$$c = 7$$

To find the length of leg a, substitute the known values into the Pythagorean Theorem.

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

$$a^2+6^2=7^2$$

Solve for a^2. Think: what number, when added to 36, gives you 49?

$$a^2+36=49$$

$$a^2=13$$

Use a calculator to find the square root of 13. The calculator gives an answer of 3.6055…, which you can round to 3.6. (Since you are approximating, you use the symbol ≈.)

$$a\approx 3.6$$

Answer: $a \approx 3.6$

9.4: Geometric Symmetry and Tessellations

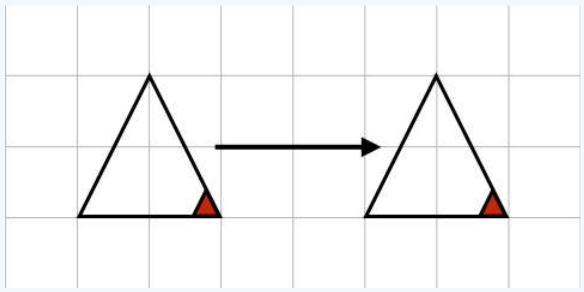
Rigid Motion

A **rigid motion** is the action of taking a geometric object in the plane and moving it is some fashion to another position in the plane without changing its shape or size. This is also called a **transformation**.

Translation

A **translation** is a transformation that moves a figure (without altering dimensions) to a new position along a line segment.

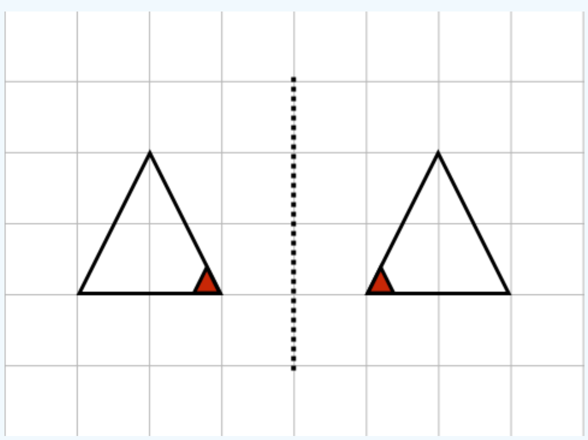
Example 1



Reflection

A **reflection** is a transformation that moves an object so that the ending position is a mirror image of the object in its starting position.

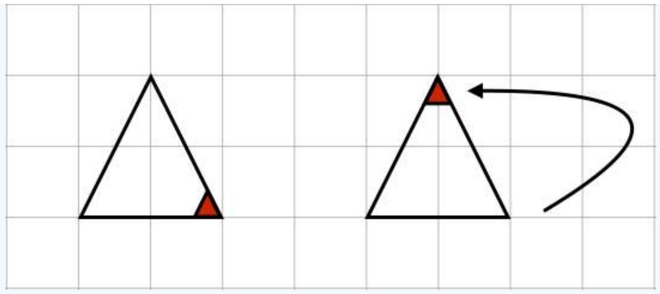
Example 2



Rotation

A **rotation** is a transformation where a figure is rotated about its center by a specified amount (given usually in degrees).

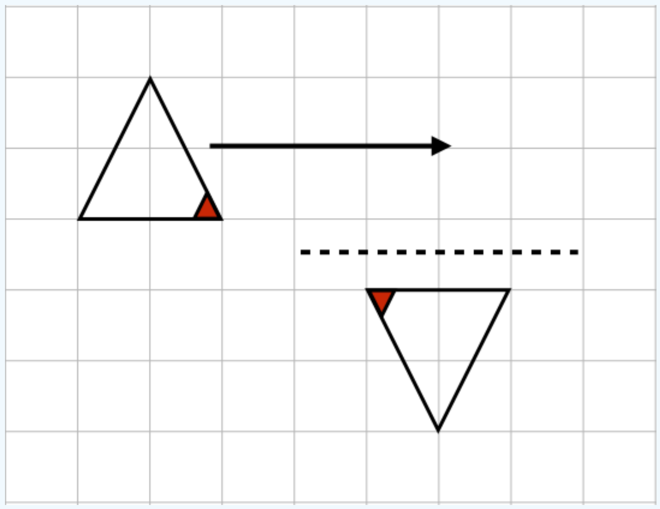
Example 3



Glide Reflection

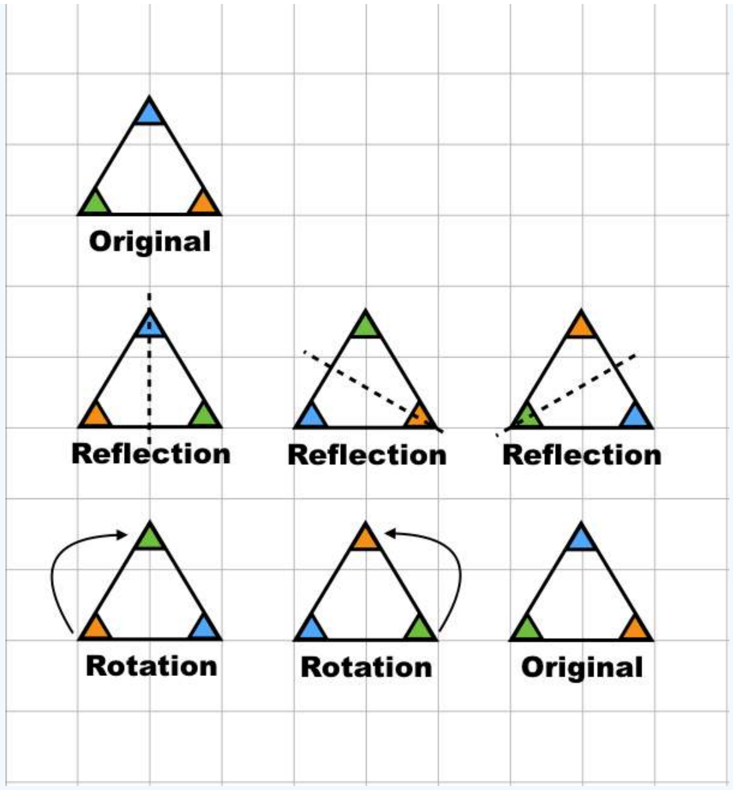
A glide reflection is a transformation formed by performing a translation followed by a reflection.

Example 4



Example 5

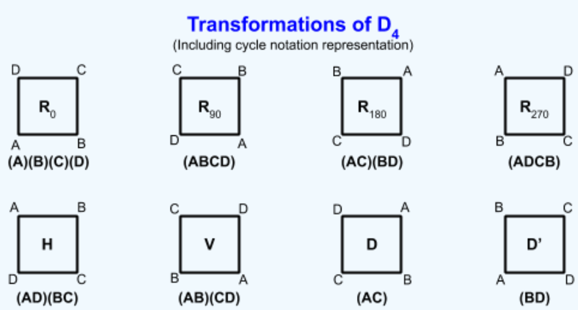
Consider the equilateral triangle below. What rigid motions can we do to it that will result in the triangle occupying the same space? What happens if we combine rigid motions?



? Try it Now 1

Consider a square. What rigid motions can we do to it that will result in the square occupying the same space? What happens if we combine rigid motions? Keep in mind that "doing nothing" is also a rigid motion.

Answer



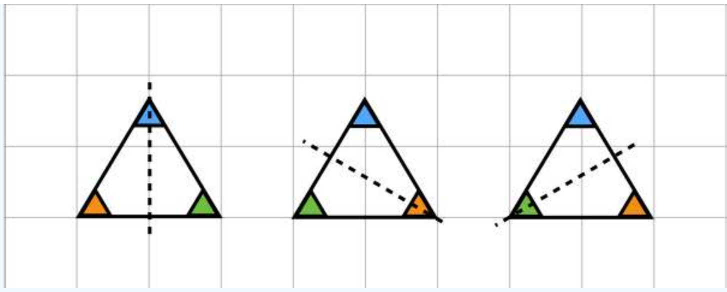
In 2D geometry, a figure is symmetrical if an operation can be done to it that leaves the figure occupying an identical physical space. This can be accomplished in two ways.

Line symmetry

Line symmetry occurs when a line may be passed through an object such that both halves of the object perfectly mirror each other.

✓ Example 6

Consider the triangle below. An equilateral triangle has three instances of line symmetry: one from each vertex to the midpoint of the opposite side.

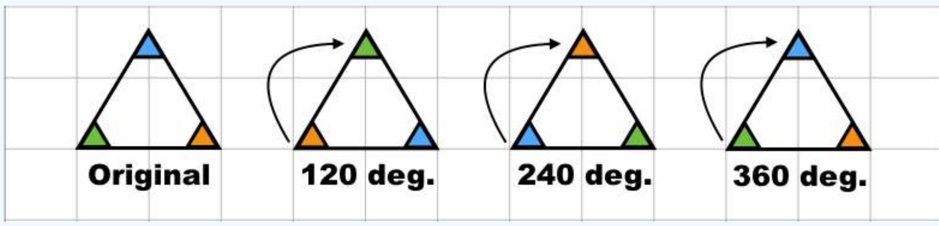


Rotational Symmetry

Rotational symmetry occurs when a shape may be rotated to occupy the same space as the original.

Example 7

Let's take the same triangle again. An equilateral triangle has three degrees of rotational symmetry: at 120, 240, and 360 degrees. Once we have determined rotational symmetry to 360 degrees, we can stop, as the pattern will repeat itself after that.



Symmetry in Art

Symmetry in art is used as a way to emphasize beauty and order. The human brain finds symmetry to be attractive and beautiful. For this reason, artists attempt to emphasize (at certain times and using certain methods) the symmetry of their subject or surroundings.

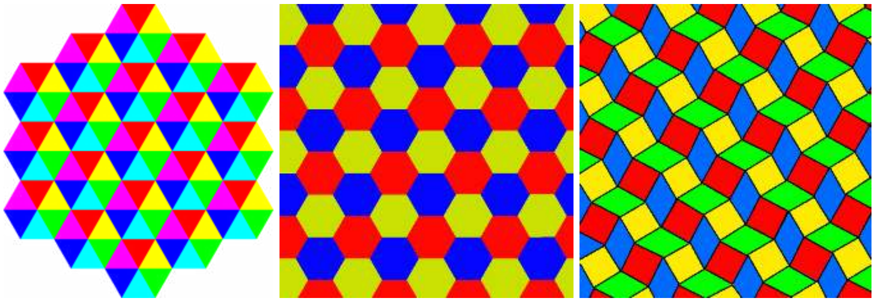
Triskelions - an example of rotational symmetry.

Tessellations

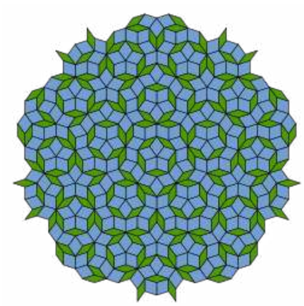
Tessellation

A **tessellation** is a design using one ore more geometric shapes with no overlaps and no gaps. The idea is that the design could be continued infinitely far to cover the whole plane (though of course we can only draw a small portion of it).

See the photos below [1] for examples.



Many tessellations have translational symmetry, but it’s not strictly necessary. The *Penrose tiling* shown below [2] does not have any translational symmetry.



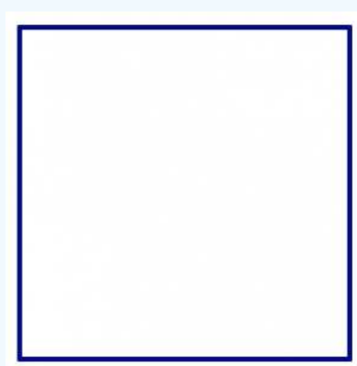
It’s actually much harder to come up with these “aperiodic” tessellations than to come up with ones that have translational symmetry. So we’ll focus on how to make symmetric tessellations.

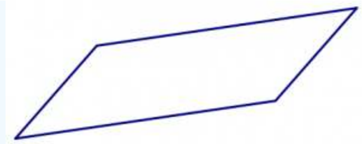
The first two tessellations above were made with a single geometric shape (called a *tile)* designed so that they can fit together without gaps or overlaps. The third design uses two basic tiles. Tessellations are often called *tilings*, and that’s what you should think about: If I had tiles made in this shape, could I use them to tile my kitchen floor? Or would it be impossible?

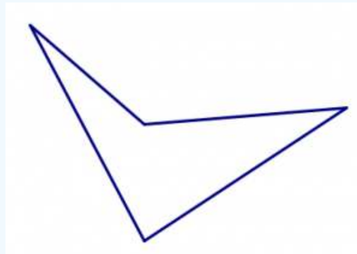
? Try it Now 2

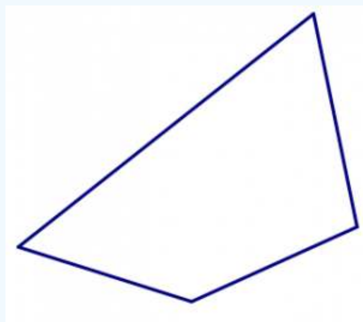
Work on these exercises on your own. You will need lots of copies (maybe 10–15 each) of each shape below. Focus on just a single tile for making your tessellation.

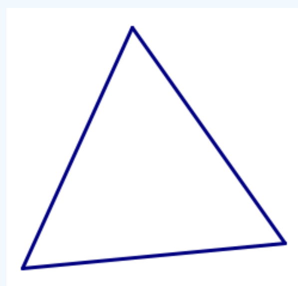
1. Start with the square tile. Can you fit the squares together in a pattern that could be continued forever, with no gaps and no overlaps? Can you do it in more than one way?
2. Now try one of the triangular tiles. Can you use many copies of a single triangle to tessellate the plane?
3. Repeat this process with each of the other tiles. Keep track of your findings.

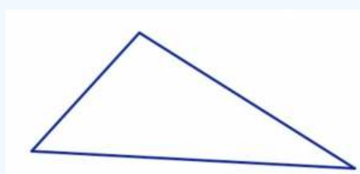


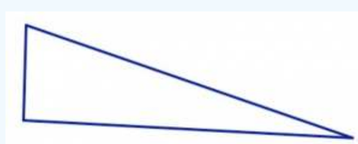


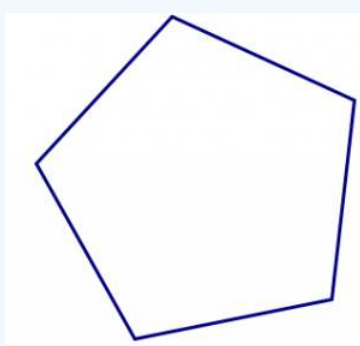


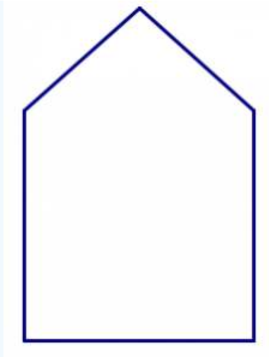


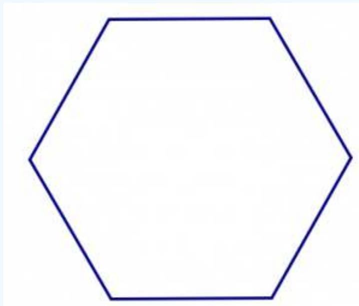


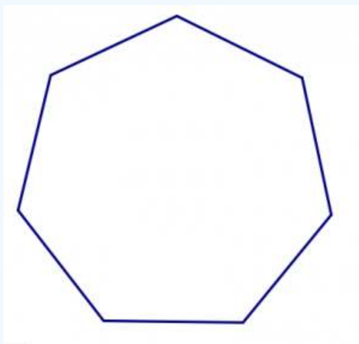


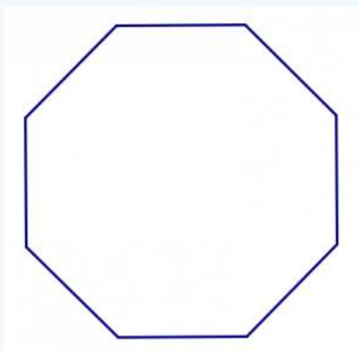












Answer

Answers may vary

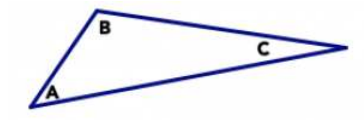
Which of the tiles given above tessellate, and which do not?

Do you have any conjectures based on this experience, about which shapes will tessellate and which will not?

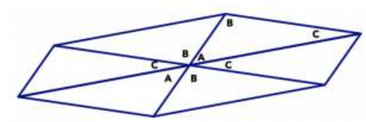
Theorem: Tessellations

Any triangle will tessellate. So will any quadrilateral.

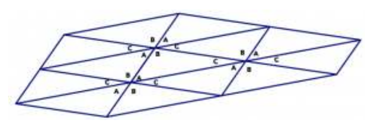
The explanation for this comes down to what you know about the sums of angles. The sum of the angles in a triangle is 180°.



So if you make six copies of a single triangle and put them together at a point so that each angle appears twice, there will be a total of 360° around the point, meaning the triangles fit together perfectly with no gaps and no overlaps.



You can then repeat this at every vertex, using more and more copies of the same triangles.



Escher Drawings

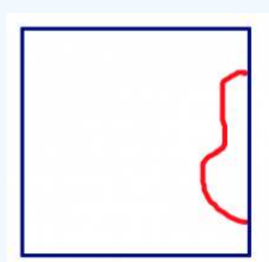
The artist M.C. Escher created many works of art inspired by mathematics, including some very beautiful tessellations. Below you will see some images [3] inspired by his work. You can view the real thing at http://www.mcescher.com/ in the "Symmetry" gallery.

You can make your own Escher-like drawings using some facts that you learned while studying tessellations.

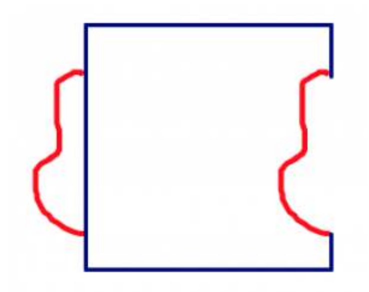
Try it Now 3

Here's how you can create your own Escher-like drawings.

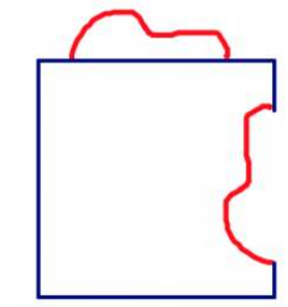
1. Select your basic tile. The first time you do this, it's easiest to start with a simple shape that you know will tessellate, like an equilateral triangle, a square, or a regular hexagon.
2. Draw a "squiggle" on one side of your basic tile.



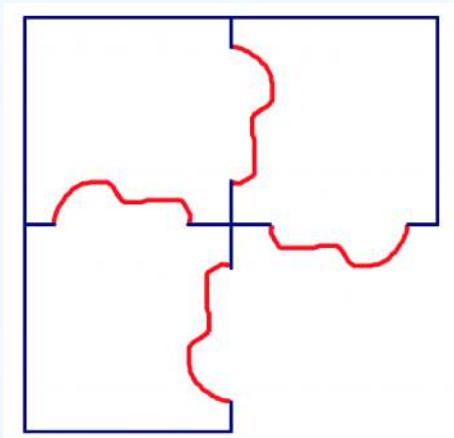
3. Cut out the squiggle, and move it to another side of your shape. You can either translate it straight across or rotate it.



or



4. It’s important that the cut-out lines up along the new edge in the same place that it appeared on its original edge.
5. Tape the squiggle into its new location. This is your basic tile. On a large piece of paper, trace around your tile. Then move it the same way you moved the squiggle (translate or rotate) so that the squiggle fits in exactly where you cut it out.



6. The shape will still tessellate, so go ahead and fill up your paper.
7. Now get creative. Color in your basic shape to look like something — an animal? a flower? a colorful blob? Add color and design throughout the tessellation to transform it into your own Escher-like drawing.
8. If you want to try a more complicated version, cut two different squiggles out of two different sides, and move them both.

Answer

Answers will vary

---

1. Triangular tessellation from pixababy [CC0]. Hexagonal and rhombic tessellations from Wikimedia Commons [Public domain]. ↵
2. Image via Wikimedia Commons [Public donmain]. ↵
3. Images from flickr [CC BY-NC-SA 2.0]. Birds by Sharon Drummond. Lizard tiles by Ben Lawson. ↵

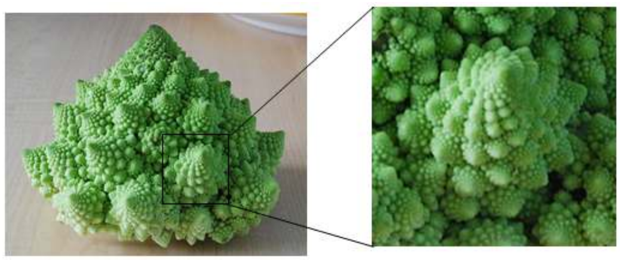
9.5: Fractals

Fractals are mathematical sets, usually obtained through recursion, that exhibit interesting dimensional properties. We'll explore what that sentence means through the rest of the section. For now, we can begin with the idea of self-similarity, a characteristic of most fractals.

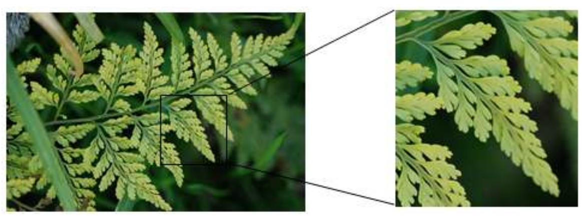
Self-Similarity

A shape is **self-similar** when it looks essentially the same from a distance as it does closer up.

Self-similarity can often be found in nature. In the Romanesco broccoli pictured below [1], if we zoom in on part of the image, the piece remaining looks similar to the whole.



Likewise, in the fern frond below [2], one piece of the frond looks similar to the whole.

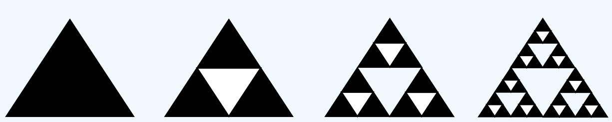


Similarly, if we zoom in on the coastline of Portugal [3], each zoom reveals previously hidden detail, and the coastline, while not identical to the view from further way, does exhibit similar characteristics.

This self-similar behavior can be replicated through recursion: repeating a process over and over.

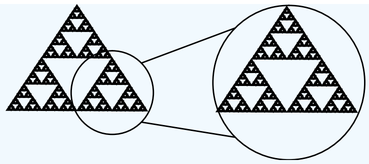
Example 1

Suppose that we start with a filled-in triangle. We connect the midpoints of each side and remove the middle triangle. We then repeat this process.



Initial Step 1 Step 2 Step 3

If we repeat this process, the shape that emerges is called the Sierpinski gasket. Notice that it exhibits self-similarity – any piece of the gasket will look identical to the whole. In fact, we can say that the Sierpinski gasket contains three copies of itself, each half as tall and wide as the original. Of course, each of those copies also contains three copies of itself.



We can construct other fractals using a similar approach. To formalize this a bit, we’re going to introduce the idea of initiators and generators.

Initiators and Generators

- An **initiator** is a starting shape
- A **generator** is an arranged collection of scaled copies of the initiator

To generate fractals from initiators and generators, we follow a simple rule:

Fractal Generation Rule

At each step, replace every copy of the initiator with a scaled copy of the generator, rotating as necessary

This process is easiest to understand through example.

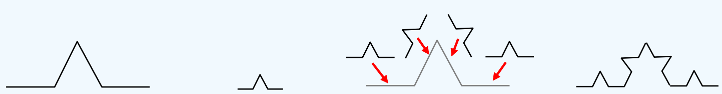
Example 2

Use the initiator and generator shown to create the iterated fractal.

initiator generator

Solution

This tells us to, at each step, replace each line segment with the spiked shape shown in the generator. Notice that the generator itself is made up of 4 copies of the initiator. In step 1, the single line segment in the initiator is replaced with the generator. For step 2, each of the four line segments of step 1 is replaced with a scaled copy of the generator:



Step 1 | Scaled copy of generator | A scaled copy replaces each line segment of Step 1 | Step 2

This process is repeated to form Step 3. Again, each line segment is replaced with a scaled copy of the generator.

Step 2 | Scaled copy of generator | Step 3

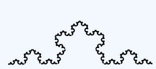
Notice that since Step 0 only had 1 line segment, Step 1 only required one copy of Step 0.

Since Step 1 had 4 line segments, Step 2 required 4 copies of the generator.

Step 2 then had 16 line segments, so Step 3 required 16 copies of the generator.

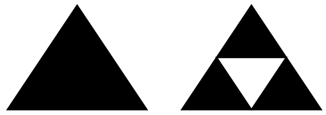
Step 4, then, would require $16 \times 4 = 64$ copies of the generator.

The shape resulting from iterating this process is called the **Koch curve**, named for Helge von Koch who first explored it in 1904.



Koch curve

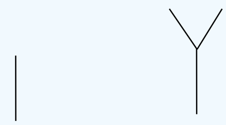
Notice that the Sierpinski gasket can also be described using the initiator-generator approach



Initiator Generator

✔ Example 3

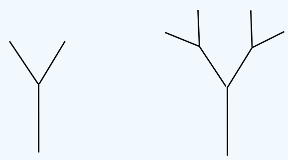
Use the initiator and generator below, however only iterate on the “branches.” Sketch several steps of the iteration.



initiator generator

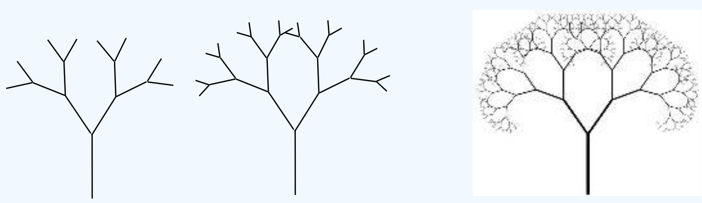
Solution

We begin by replacing the initiator with the generator. We then replace each “branch” of Step 1 with a scaled copy of the generator to create Step 2.



Step 1 Step 2

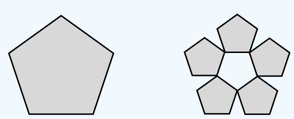
We can repeat this process to create later steps. Repeating this process can create intricate tree shapes [4].



Step 3 Step 4 Final shape

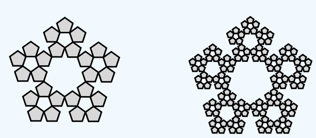
? Try it Now 1

Use the initiator and generator shown to produce the next two stages



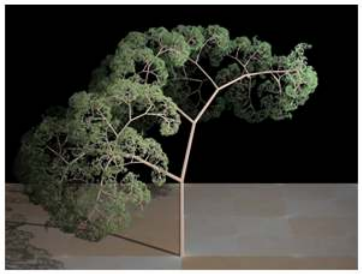
Initiator Generator

Answer

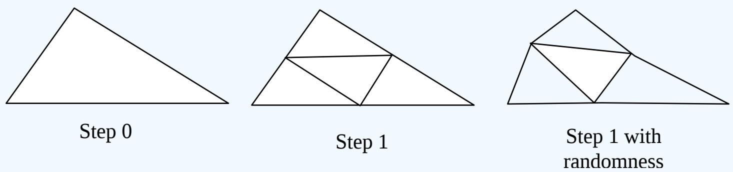
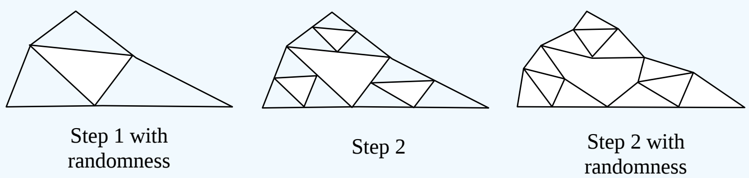


Step 2 Step 3

Using iteration processes like those above can create a variety of beautiful images evocative of nature [5][6].



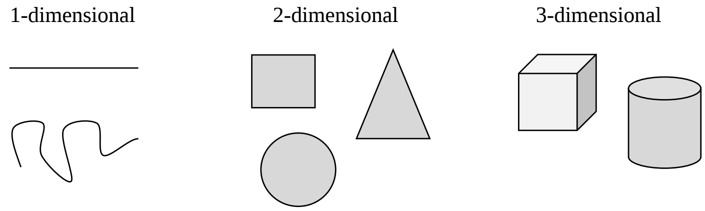
More natural shapes can be created by adding in randomness to the steps.

> **✔ Example 4**
>
> Create a variation on the Sierpinski gasket by randomly skewing the corner points each time an iteration is made.
>
> **Solution**
>
> Suppose we start with the triangle below. We begin, as before, by removing the middle triangle. We then add in some randomness.
>
> 
>
>
> We then repeat this process.
>
> 
>
>
> Continuing this process can create mountain-like structures.

The landscape [7] below was created using fractals, then colored and textured.

In addition to visual self-similarity, fractals exhibit other interesting properties. For example, notice that each step of the Sierpinski gasket iteration removes one quarter of the remaining area. If this process is continued indefinitely, we would end up essentially removing all the area, meaning we started with a 2-dimensional area, and somehow end up with something less than that, but seemingly more than just a 1-dimensional line.

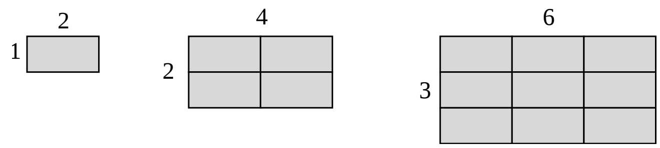
To explore this idea, we need to discuss dimension. Something like a line is 1-dimensional; it only has length. Any curve is 1-dimensional. Things like boxes and circles are 2-dimensional, since they have length and width, describing an area. Objects like boxes and cylinders have length, width, and height, describing a volume, and are 3-dimensional.



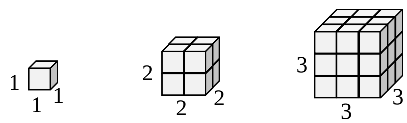
Certain rules apply for scaling objects, related to their dimension.

If I had a line with length 1, and wanted scale its length by 2, I would need two copies of the original line. If I had a line of length 1, and wanted to scale its length by 3, I would need three copies of the original.

If I had a rectangle with length 2 and height 1, and wanted to scale its length and width by 2, I would need four copies of the original rectangle. If I wanted to scale the length and width by 3, I would need nine copies of the original rectangle.



If I had a cubical box with sides of length 1, and wanted to scale its length, its width, and its height by 2, I would need eight copies of the original cube. If I wanted to scale the length, width, and height by 3, I would need 27 copies of the original cube.



Notice that in the 1-dimensional case, copies needed = scale.

- In the 2-dimensional case, copies needed $=$ scale^2.
- In the 3-dimensional case, copies needed $=$ scale^3.

From these examples, we might infer a pattern.

Scaling-Dimension Relation

To scale a D-dimensional shape by a scaling factor S, the number of copies C of the original shape needed will be given by:

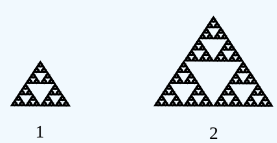
Copies $=$ Scale $^{\text{Dimension}}$, or $C = S^D$

Example 5

Use the scaling-dimension relation to determine the dimension of the Sierpinski gasket.

Solution

Suppose we define the original gasket to have side length 1. The larger gasket shown is twice as wide and twice as tall, so has been scaled by a factor of 2.



Notice that to construct the larger gasket, 3 copies of the original gasket are needed.

Using the scaling-dimension relation $C = S^D$, we obtain the equation $3 = 2^D$

since $2^1 = 2$ and $2^2 = 4$, we can immediately see that D is somewhere between 1 and 2; the gasket is more than a 1 - dimensional shape, but we've taken away so much area its now less than 2-dimensional.

Solving the equation $3 = 2^D$ requires logarithms. If you studied logarithms earlier, you may recall how to solve this equation (if not, just skip to the box below and use that formula):

$$\begin{array}{ll} 3 = 2^D & \text{Take the logarithm of both sides} \\ \log(3) = \log\left(2^D\right) & \text{Use the exponent property of logs} \\ \log(3) = D\log(2) & \text{Divide by } \log(2) \\ D = \frac{\log(3)}{\log(2)} \approx 1.585 & \text{The dimension of the gasket is about } 1.585 \end{array} \tag{9.5.1}$$

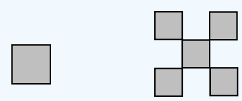
Scaling-Dimenstion Relation To Find Dimenstion

To find the dimension D of a fractal, determine the scaling factor S and the number of copies C of the original shape needed, then use the formula

$$D = \frac{\log(C)}{\log(S)}$$

Try it Now 2

Determine the fractal dimension of the fractal produced using the initiator and generator



Initiator Generator

Answer

1 3

Scaling the fractal by a factor of 3 requires 5 copies of the original. $D = \frac{\log(5)}{\log(3)} \approx 1.465$

[1] en.Wikipedia.org/wiki/File:Ca...ractal_AVM.JPG

[2] http://www.flickr.com/photos/cjewel/3261398909/

[3] Openstreetmap.org, CC-BY-SA

[4] http://www.flickr.com/photos/visualarts/5436068969/

[5] en.Wikipedia.org/wiki/File:Fr...e_b_-_2%29.jpg

[6] en.Wikipedia.org/wiki/File:Ba...-_4_states.PNG

[7] en.Wikipedia.org/wiki/File:Fr...lLandscape.jpg

9.6: Sequences

Sequence

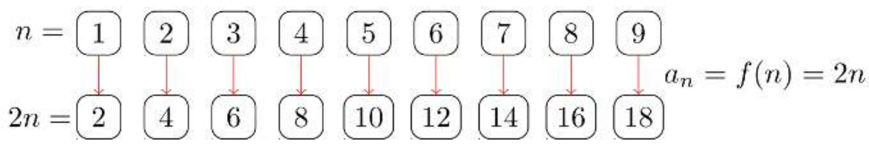
A **sequence** is a list of numbers: $a_1, a_2, a_3, a_4, \ldots, a_n, \ldots$ A sequence can be a finite or infinite list. We call a_1 the first term, a_2 the second term, and a_n the "general term" or the n^{th} term. Sequences have a pattern. We describe the pattern in the general term a_n.

The following sequence of numbers has a pattern you are bound to recognize:

$$2, 4, 6, 8, 10, 12, 14, 16, 18, \ldots$$

Likely, you would describe the sequence in words: *the sequence of even numbers.* Alternatively, can we describe the sequence mathematically? That is, can we describe the pattern of the sequence of even numbers using a formula? Absolutely! This section will explore arithmetic sequences, how to identify them, mathematically describe their terms, and the relationship between arithmetic sequences and linear functions. Let's get started!

For the sequence of even numbers: $2, 4, 6, 8, 10, \ldots$ the general term $a_n = 2n$.



The general term a_n of a sequence is simply a function of n, indicated above as $f(n)$, where n is a natural number.

Example 1

In the sequence of even numbers, what is the 20^{th} term in the sequence?

Solution

The general term of the sequence of even numbers is $a_n = 2n$. Since $n =$ the term number, we are asked to find a_{20}.

Plug in the term-number $n = 20$ into the formula $a_n = 2n$

$a_{20} = 2(20) = 40$

Arithmetic Sequence

If the sequence: $a_1, a_2, a_3, a_4, \ldots, a_{n-1}, a_n, \ldots$ exhibits a pattern such that

$$a_n - a_{n-1} = d \tag{9.6.1}$$

For all n, then the real number d is called the **common difference**, and the sequence is an **arithmetic sequence**.

Example 2

A sequence is given. If the sequence is an arithmetic sequence, give the common difference. If the sequence is not an arithmetic sequence, explain how it fails to be arithmetic.

a. $25, 32, 39, 46, 53, 60, \ldots$
b. $2, 4, 8, 16, 32, \ldots$
c. $3^2, 3^4, 3^6, 3^8, 3^{10}, \ldots$
d. $0, 1, 0, 1, 0, 1, \ldots$

Solution

a. Is the sequence

$$25, 32, 39, 46, 53, 60, \ldots$$

an arithmetic sequence?

$$\begin{aligned} a_2-a_1 &= 32-25 &= 7 \\ a_3-a_2 &= 39-32 &= 7 \\ a_4-a_3 &= 46-39 &= 7 \\ a_5-a_4 &= 53-46 &= 7 \\ a_6-a_5 &= 60-53 &= 7 \end{aligned}$$

The sequence is arithmetic and the common difference is 7.

b. Is the sequence

$$2, 4, 8, 16, 32, \ldots$$

an arithmetic sequence?

$$\begin{aligned} a_2-a_1 &= 4-2 &= 2 \\ a_3-a_2 &= 8-4 &= 4 \\ a_4-a_3 &= 16-8 &= 8 \\ a_5-a_4 &= 32-16 &= 16 \\ & & 2 \neq 4 \neq 8 \neq 16 \end{aligned}$$

The sequence is not arithmetic. $a_n - a_{n-1}$ does not yield a common difference.

c. Is the sequence

$$3^2, 3^4, 3^6, 3^8, 3^{10}, \ldots$$

an arithmetic sequence?

$$\begin{aligned} 3^4-3^2 &= 3^2(3^2-1) &= 9\cdot 8 &= 72 \\ 3^6-3^4 &= 3^4(3^2-1) &= 81\cdot 8 &= 648 \end{aligned}$$

Since $a_3 - a_2 \neq a_2 - a_1$, we conclude the sequence is not arithmetic.

d. Is the sequence

$$0, 1, 0, 1, 0, 1, \ldots$$

an arithmetic sequence?

$$\begin{aligned} -0 &= 1 \\ 0-1 &= -1 \end{aligned}$$

Since $a_3 - a_2 \neq a_2 - a_1$, the sequence is not arithmetic.

If a sequence is arithmetic, the general term a_n is determined using the common difference, d, of the sequence. Functions of the form $y = mx + b$, known as linear functions, have a strong relationship to arithmetic sequences. The slope m of a linear function is equivalent to the common difference d of an arithmetic sequence. Let's compare arithmetic sequences to linear functions to build a_n, the general term of an arithmetic sequence.

✓ Example 3

Find the general term a_n of each arithmetic sequence:

a. $4, 7, 10, 13, 16, \ldots$
b. $100, 80, 60, 40, 20, \ldots$

Solution

We will create a table of values for each sequence. The first column will be the term number, n, starting with $n = 1$. The second column will list the terms of the sequence. The common difference is shown on the side of the second column.

a. The sequence $4, 7, 10, 13, 16, \ldots$ has the common difference $d = 3$. But it's also the slope m of the linear function $f(x) = mx + b$.

$$m=\frac{a_n-a_{n-1}}{n-(n-1))}=\frac{d}{1}=d \tag{9.6.2}$$

n	a_n
1	4
2	7
3	10
4	13
5	16

+3 +3 +3 +3

The above table essentially mimics any linear function, $f(x)=mx+b$.

- Instead of x, sequences use $n-$values.
- Instead of $m=$ slope in linear functions, sequences use $d=$ common difference.
- Instead of b, a sequence notates the same value with a_0.

If a_1 denotes the first term of a sequence, then the general term of a sequence is:

$$a_n=f(n)=d\cdot n+a_0 \tag{9.6.3}$$

To find the general term, a_n, we will need to find the value a_0. There are several ways to do this, but perhaps the simplest is to create an extra row where $n=0$, then use the common difference to find a_0. The common difference pattern is maintained and $a_0+d=a_1$.

n	a_n
0	$a_0=?$
1	4
2	7
3	10
4	13
5	16

+3 +3 +3 +3 +3 $d=3$

Find the value a_0:

$$\begin{aligned} a_0+3 &= 4 \\ a_0+3-3 &= 4-3 \\ a_0 &= 1 \end{aligned}$$

The general term of the sequence is:

$$a_n=3n+1$$

b. Use the same strategy for Example $8.1.3$a to solve Example $8.1.3$b. Create a table, find the common difference, d, and find the a_0 term of the sequence $100, 80, 60, 40, 20, \ldots$

n	a_n
0	$a_0=?$
1	100
2	80
3	60
4	40
5	20

-20 -20 -20 -20 -20 $d=-20$

The common difference $d=-20$. Find the value a_0.

$$\begin{aligned} a_0-20 &= 100 \\ a_0-20+20 &= 100+20 \\ a_0 &= 120 \end{aligned}$$

The general term of the sequence is:

$$a_n = -20n + 120$$

The General n^{th} Term of an Arithmetic Sequence

An **arithmetic sequence** with first term a_1 and common difference d is given by $a_n = a_1 + (n-1)d, \quad n \geq 1$ for n^{th} term.

Example 4

A five-year old child receives an allowance of $\$1$ each week. Their parents promise them an annual increase of $\$2$ per week.

a. Write a formula for the child's weekly allowance in a given year.
b. What will the child's allowance be when he is 16 years old?

Solution

a. The situation can be modeled by an arithmetic sequence with an initial term of 1 and a common difference of 2.

Let A be the amount of the allowance and n be the number of years after age 5. Using the altered explicit formula for an arithmetic sequence we get:

$$A_n = 1 + 2n$$

b. We can find the number of years since age 5 by subtracting.

$$16 - 5 = 11$$

We are looking for the child's allowance after 11 years. Substitute 11 into the formula to find the child's allowance at age 16.

$$A_{11} = 1 + 2(11) = 23$$

The child's allowance at age 16 will be $\$23$ per week.

Geometric sequences, on the other hand, have a common ratio. Each term after the first term is obtained by multiplying the previous term by r, the common ratio. As an example, the following sequence does not have a common difference, so it is not an arithmetic sequence. Instead, this sequence has a **common ratio**, r:

$$\frac{a_n}{a_{n-1}} = r = 2$$

and it is a geometric sequence:

$$2, 4, 8, 16, 32, 64, 128$$

Notice that each term is double the previous term. By multiplying any term by 2, we obtain the subsequent term. A common ratio is the hallmark signature of a geometric sequence.

Geometric Sequence

If the sequence: $a_1, a_2, a_3, a_4, \ldots, a_{n-1}, a_n, \ldots$ exhibits a pattern ($a \neq 0$ and $r \neq 0$) such that

$a_1, a_1 r, a_1 r^2, a_1 r^3, \ldots, a_1 r^{n-1}, a_1 r^n, \ldots$

Then the sequence is geometric and r is called the common ratio, where $\frac{a_n}{a_{n-1}} = r$

A geometric sequence is analogous to an exponential function, $f(x) = ab^x$, where a and b are constants, $a =$ any real number and $b > 0$. The general term a_n for a geometric sequence will mimic the exponential function formula, but modified in the following way:

- Instead of $x =$ any real number, the domain of the geometric sequence function is the set of natural numbers n.
- The constant a will become the first term, or a_1, of the geometric sequence.
- The constant b is replaced by the common ratio r, but r can be positive or negative.

The General Term of a Geometric Sequence

The geometric sequence with common ratio r:

$a_1, a_1 r, a_1 r^2, a_1 r^3, \ldots, a_1 r^{n-1}, \ldots$

has general term

$a_n = a_1 r^{n-1}$

The first term of the geometric sequence is a_1, or $a_1 r^0$. Recall that $r^0 = 1$.It's worth mentioning that, in some cases, the first term is better notated as a_0 rather than a_1. If we use $a_0 =$ first term (starting the sequence at $n = 0$), then the geometric sequence would be notated: $a_0, a_0 r, a_0 r^2, a_0 r^3$, … and the general term is $a_n = a_0 r^n$. Although the term-number no longer matches the subscript (i.e. $a_1 =$ second term, $a_2 =$ third term, etc.), the exponent on r tells us how many times r was applied. In real-life problems that have an initial value to which r is repetitively multiplied, allow yourself the flexibility to call the initial amount a_0.

Example 5

Determine the common ratio of the geometric sequence: $15, 45, 135, 405, ..$ and give the general term, a_n. Then find the 10^{th} term of the sequence, or a_{10}.

Solution

Finding the common ratio is a matter of dividing any term by its previous term:

$$\frac{45}{15} = 3 = r\,.$$

Therefore, the general term of the sequence is:

$$a_n = 15 \cdot 3^{n-1}$$

The general term gives us a formula to find a_{10}. Plug $n = 10$ into the general term a_n.

$$a_{10} = 15 \cdot 3^{10-1} = 15 \cdot 3^9 = 295245$$

Example 6

Determine the common ratio of the geometric sequence: $8, -12, 18, -27, \ldots$and give the general term a_n. Then find the 7^{th} term of the sequence.

Solution

Notice the sequence alternates in sign value: positive, negative, positive, negative, … An alternating sequence occurs when $r < 0$. We expect the r-value to be negative.

The common ratio is found by dividing two consecutive terms. Let's divide $\frac{a_2}{a_1}$.

$$\frac{-12}{8} = -\frac{3}{2} = -1.5 = r$$

Therefore, the general term of the sequence is:

$$a_n = -1.5 \cdot 8^{n-1}$$

The general term gives us a formula to find a_7. Plug into $n = 7$ in a_n to find the 7^{th} term:

$$a_7 = 8(-1.5)^{7-1} = 8(-1.5)^6 = 91.125$$

Example 7

A filtering process can reduce Chemical B by 10%. The process can be repeated and have the same reduction rate of 10% each time. Initially, there is 7 mg of Chemical B before filtering. How much of Chemical B remains after 4 filtering processes? Round the answer to 2 decimal places.

Solution

If chemical B is reduced by 10%, then 90% remains after filtering. The sequence would end after 4 filters. The common ratio $r = 0.9$ and we apply this common ratio 4 times to the initial value, 7 mg:

Rather than performing each multiplication separately, it's easier to compute the remaining quantity of Chemical B using the formula for the general term:

$$7 \cdot (0.9)^4 = 4.5927$$

Answer After 4 filtering processes, 4.59 mg of Chemical B remains.

A famous and important sequence is the Fibonacci sequence, named after the Italian mathematician known as Leonardo Pisano, whose nickname was Fibonacci, and who lived from 1170 to 1230. This sequence is:

$$\{1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots\ldots\ldots\} \tag{9.6.4}$$

This sequence is defined recursively. This means each term is defined by the previous terms.

Fibonacci Sequence

The Fibonacci sequence is defined by $f_1 = 1, f_2 = 1, f_n = f_{n-1} + f_{n-2}$.

In other words, to get the next term in the sequence, add the two previous terms.

$\{1, 1, 1+1 = 2, 2+1 = 3, 3+2 = 5, 5+3 = 8, 8+5 = 13, 13+8 = 21, 21+13 = 34, 34+21 = 55, 55+34 = 89, 89+55 = 144, \cdots\}$

Example 8

Find the 13th, 14th, and 15th Fibonacci numbers

Solution

First, notice that there are already 12 Fibonacci numbers listed above, so to find the next three Fibonacci numbers, we simply add the two previous terms to get the next term as the definition states.

$f_{13} = 144 + 89 = 233, f_{14} = 233 + 144 = 377, f_{15} = 377 + 233 = 610$

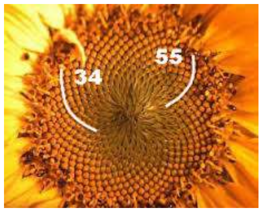
We can find the Fibonacci numbers in all around us in nature. The number of branches on some trees or the number of petals of some daisies are often Fibonacci numbers

a. Daisy with 13 petals b. Daisy with 21 petals

a.  b. 

Fibonacci numbers also appear in spiral growth patterns such as the number of spirals on a cactus or in sunflowers seed beds.

a. Cactus with 13 clockwise spirals b. Sunflower with 34 clockwise spirals and 55 counterclockwise spirals

a.  b. 

Another interesting fact arises when looking at the ratios of consecutive Fibonacci numbers.

Table 9.6.1:

Fibonacci number	divided by the one before	ratio
1		
1	1/1	= 1.0000
2	2/1	= 2.0000
3	3/2	= 1.5000
5	5/3	= 1.6667
8	8/5	= 1.6000
13	13/8	= 1.6250
21	21/13	= 1.6154...
34	34/21	= 1.6190...
55	55/34	= 1.6177...
89	89/55	= 1.6182...
...	...	...
		= 1.6180...

The number that these ratios are getting closer to is a special number called the Golden Ratio which is denoted by ϕ (the Greek letter phi).

The Golden Ratio

The Golden Ratio: $\phi = \frac{1+\sqrt{5}}{2}$. The Golden Ratio has the decimal approximation of $\phi = 1.6180339887$.

The Golden Ratio is a special number for a variety of reasons. It is also called the divine proportion and it appears in art and architecture. It is claimed by some to be the most pleasing ratio to the eye.

The Golden Rectangle

If the length of a rectangle divided by its width is equal to the Golden Ratio, then the rectangle is called a **golden rectangle**.

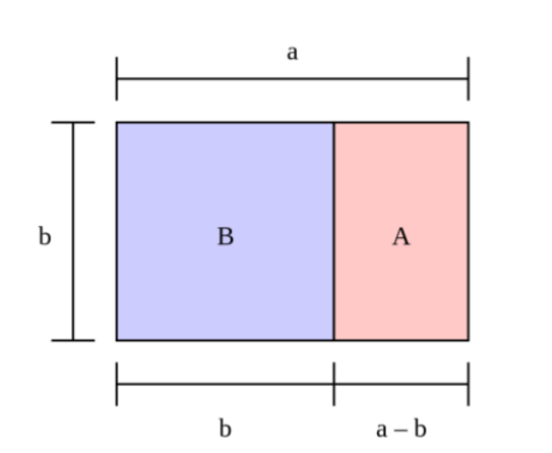
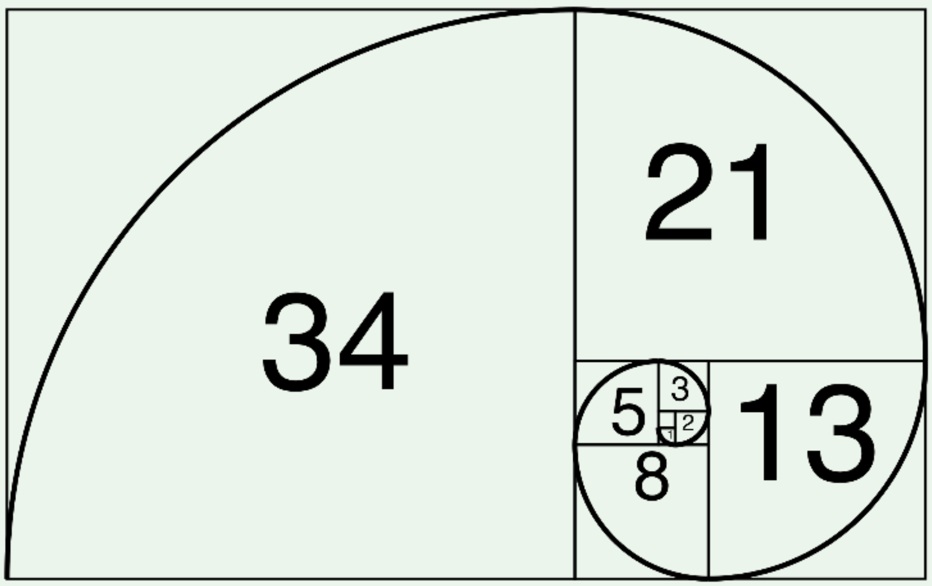
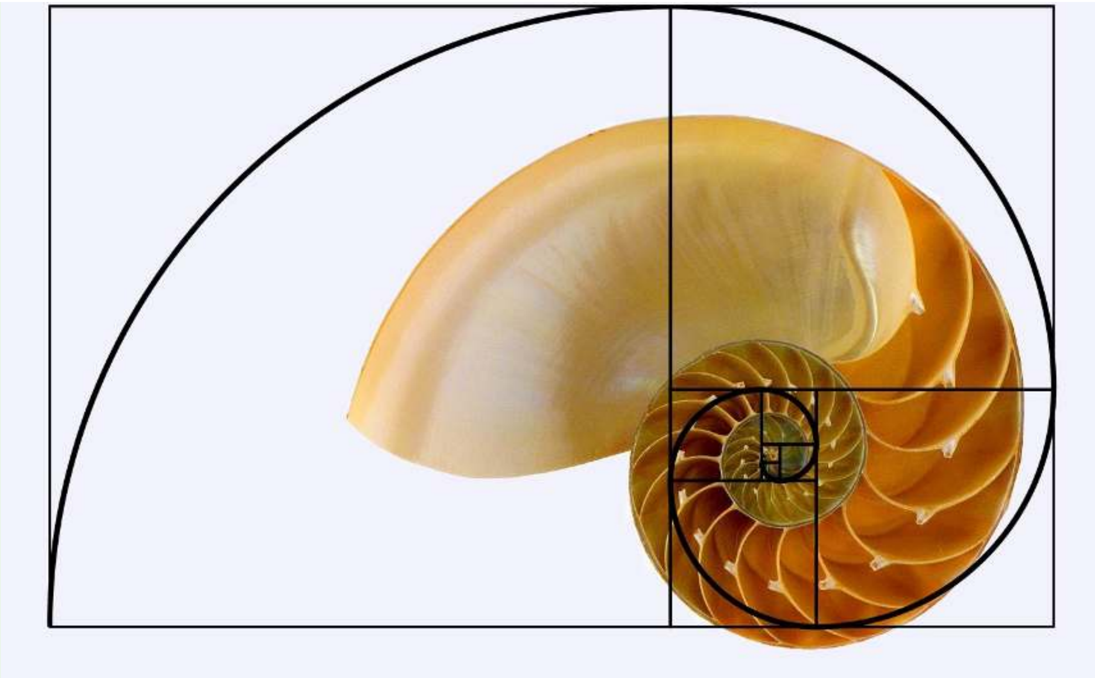


Figure 9.6.1: Image by Peter John Acklam is licensed by CC-3.0

If you cut off a square on one end of the rectangle, the smaller rectangle is again a golden rectangle. If you continue doing this, and connect the opposite diagonals of the squares that you are cutting off with a smooth curve, you generate a spiral.



Here is a nautilus shell and the golden rectangle spiral on top.



CHAPTER OVERVIEW

10: Apportionment

10.1: Hamilton's Method

Apportionment is the problem of dividing up a fixed number of things among groups of different sizes. In politics, this takes the form of allocating a limited number of representatives amongst voters. This problem, presumably, is older than the United States, but the best-known ways to solve it have their origins in the problem of assigning each state an appropriate number of representatives in the new Congress when the country was formed. States also face this apportionment problem in defining how to draw districts for state representatives. The apportionment problem comes up in a variety of non-political areas too, though. We face several restrictions in this process:

Apportionment Rules

1. The things being divided up can exist only in whole numbers.
2. We must use all of the things being divided up, and we cannot use any more.
3. Each group must get at least one of the things being divided up.
4. The number of things assigned to each group should be at least approximately proportional to the population of the group. (Exact proportionality isn't possible because of the whole number requirement, but we should try to be close, and in any case, if Group A is larger than Group B, then Group B shouldn't get more of the things than Group A does.)

In terms of the apportionment of the United States House of Representatives, these rules imply:

1. We can only have whole representatives (a state can't have 3.4 representatives)
2. We can only use the (currently) 435 representatives available. If one state gets another representative, another state has to lose one.
3. Every state gets at least one representative
4. The number of representatives each state gets should be approximately proportional to the state population. This way, the number of constituents each representative has should be approximately equal.

Alexander Hamilton proposed the method that now bears his name. His method was approved by Congress in 1791, but was vetoed by President Washington. It was later adopted in 1852 and used through 1911. He begins by determining, to several decimal places, how many things each group should get. Since he was interested in the question of Congressional representation, we'll use the language of states and representatives, so he determines how many representatives each state should get. He follows these steps:

Hamilton's Method

1. Determine how many people each representative should represent. Do this by dividing the total population of all the states by the total number of representatives. This answer is called the **divisor**.
2. Divide each state's population by the divisor to determine how many representatives it should have. Record this answer to several decimal places. This answer is called the **quota**. Since we can only allocate whole representatives, Hamilton resolves the whole number problem, as follows:
3. Cut off all the decimal parts of all the quotas (but don't forget what the decimals were). These are called the **lower quotas**. Add up the remaining whole numbers. This answer will always be less than or equal to the total number of representatives (and the "or equal to" part happens only in very specific circumstances that are incredibly unlikely to turn up).
4. Assuming that the total from Step 3 was less than the total number of representatives, assign the remaining representatives, one each, to the states whose decimal parts of the quota were largest, until the desired total is reached.

Make sure that each state ends up with at least one representative!

Note on rounding: Today we have technological advantages that Hamilton (and the others) couldn't even have imagined. Take advantage of them, and keep several decimal places.

Example 1

The state of Delaware has three counties: Kent, New Castle, and Sussex. The Delaware state House of Representatives has 41 members. If Delaware wants to divide this representation along county lines (which is *not* required, but let's pretend they do), let's use Hamilton's method to apportion them. The populations of the counties are as follows (from the 2010 Census):

County	Population
Kent	162,310
New Castle	538,479
Sussex	197,145
Total	**897,934**

Solution

1. First, we determine the divisor: $897,934/41 = 21,900.82927$

2. Now we determine each county's quota by dividing the county's population by the divisor:

County	Population	Quota
Kent	162,310	7.4111
New Castle	538,479	24.5872
Sussex	197,145	9.0017
Total	**897,934**	

3. Removing the decimal parts of the quotas gives:

County	Population	Quota	Initial
Kent	162,310	7.4111	7
New Castle	538,479	24.5872	24
Sussex	197,145	9.0017	9
Total	**897,934**		**40**

4. We need 41 representatives and this only gives 40. The remaining one goes to the county with the largest decimal part, which is New Castle:

County	Population	Quota	Initial	Final
Kent	162,310	7.4111	7	7
New Castle	538,479	24.5872	24	25
Sussex	197,145	9.0017	9	9
Total	**897,934**		**40**	**41**

Example 2

Use Hamilton's method to apportion the 75 seats of Rhode Island's House of Representatives among its five counties.

County	Population
Bristol	49,875
Kent	166,158
Newport	82,888
Providence	626,667
Washington	126,979
Total	**1,052,567**

Solution

1. The divisor is $1,052,567/75 = 14,034.22667$

2. Determine each county's quota by dividing its population by the divisor:

County	Population	Quota
Bristol	49,875	3.5538
Kent	166,158	11.8395
Newport	82,888	5.9061
Providence	626,667	44.6528
Washington	126,979	9.0478
Total	**1,052,567**	

3. Remove the decimal part of each quota:

County	Population	Quota	Initial
Bristol	49,875	3.5538	3
Kent	166,158	11.8395	11
Newport	82,888	5.9061	5
Providence	626,667	44.6528	44
Washington	126,979	9.0478	9
Total	**1,052,567**		**72**

4. We need 75 representatives and we only have 72, so we assign the remaining three, one each, to the three counties with the largest decimal parts, which are Newport, Kent, and Providence:

County	Population	Quota	Initial	Final
Bristol	49,875	3.5538	3	3
Kent	166,158	11.8395	11	12
Newport	82,888	5.9061	5	6
Providence	626,667	44.6528	44	45
Washington	126,979	9.0478	9	9
Total	**1,052,567**		**72**	**75**

Note that even though Bristol County's decimal part is greater than .5, it isn't big enough to get an additional representative, because three other counties have greater decimal parts.

Controversy

After seeing Hamilton's method, many people find that it makes sense, it's not that difficult to use (or, at least, the difficulty comes from the numbers that are involved and the amount of computation that's needed, not from the method), and they wonder why anyone would want another method. The problem is that Hamilton's method is subject to several paradoxes. Three of them happened, on separate occasions, when Hamilton's method was used to apportion the United States House of Representatives.

The **Alabama Paradox** is named for an incident that happened during the apportionment that took place after the 1880 census. (A similar incident happened ten years earlier involving the state of Rhode Island, but the paradox is named after Alabama.) The post-1880 apportionment had been completed, using Hamilton's method and the new population numbers from the census. Then it was decided that because of the country's growing population, the House of Representatives should be made larger. That meant that the apportionment would need to be done again, still using Hamilton's method and the same 1880 census numbers, but with more representatives. The assumption was that some states would gain another representative and others would stay with the same number they already had (since there weren't enough new representatives being added to give one more to every state). The paradox is that Alabama ended up *losing* a representative in the process, even though no populations were changed and the total number of representatives increased.

The **New States Paradox** happened when Oklahoma became a state in 1907. Oklahoma had enough population to qualify for five representatives in Congress. Those five representatives would need to come from somewhere, though, so five states, presumably, would lose one representative each. That happened, but another thing also happened: Maine gained a representative (from New York).

The **Population Paradox** happened between the apportionments after the census of 1900 and of 1910. In those ten years, Virginia's population grew at an average annual rate of 1.07%, while Maine's grew at an average annual rate of 0.67%. Virginia started with more people, grew at a faster rate, grew by more people, and ended up with more people than Maine. By itself, that doesn't mean that Virginia should gain representatives or Maine shouldn't, because there are lots of other states involved. But Virginia ended up losing a representative *to Maine*.

10.2: Huntington-Hill Method

The Huntington-Hill Method attempts to minimize the percent differences (unfairness) of how many people each representative will represent.

Huntington-Hill Method

If states X and Y have already been allotted x and y representatives respectively, then state X should be given an additional representative over state Y if:

$$\frac{(\text{population of Y})^2}{y \times (y+1)} < \frac{(\text{population of X})^2}{x \times (x+1)} \tag{10.2.1}$$

$$\frac{(\text{population of X})^2}{x \times (x+1)} \tag{10.2.2}$$

is called State X's **Huntington-Hill Number**.

Huntington-Hill Method Procedure

1. Each state gets one representative to begin with. Now calculate the Huntington-Hill number for each state with a Current Representation value of 1.
2. Create a table of Huntington-Hill numbers for each state.
3. Give the next representative to the state with the highest Huntington-Hill number.
4. Recalculate the Huntington-Hill number for that state with the next Current Representation value.
5. Repeat steps 3 and 4 until all the representatives have been given out.
6. State the results in a complete sentence.

Example 1

In a hypothetical world, State X has 750 people, State Y has 300 people, and State Z has 390 people. Use Huntington-Hill Method to apportion 9 representatives.

Solution

Step 1) State X, Y, and Z all get one representative. Note: The denominator will be 1(2).

State X has a population of 750 and 1 representative so the calculation for the Huntington-Hill number for State X is:

$$\frac{750^2}{1 \times 2} \tag{10.2.3}$$

You can calculate the Huntington-Hill number for State Y and State Z by replacing the 750 with 300 or 390 respectively.

Step 2) Create a table where X, Y, and Z have one representative.

Current Representation	State X	State Y	State Z
1	281,250	45,000	76,050

Step 3) Give the 4th representative to X because it is the highest Huntington-Hill number. Recalculate State X's Huntington-Hill number for Current Representation of 2.

Step 4) Recalculate the Huntington-Hill number for X with 2 representatives. Note: The denominator will now be 2(3)

Current Representation	State X	State Y	State Z
1	281,250	45,000	76,050
2	93,750		

Step 5) Repeat steps 3 and 4 until all 9 representatives have been assigned.

Give the 5th representative to X because it is the highest Huntington-Hill number. Recalculate State X's Huntington-Hill number for Current Representation of 3.

Current Representation	State X	State Y	State Z
1	$281,250$	$45,000$	$76,050$
2	$93,750$		
3	$46,875$		

Give the 6th representative to Z and recalculate with Current Representation of 2.

Current Representation	State X	State Y	State Z
1	$281,250$	$45,000$	$76,050$
2	$93,750$		$25,350$
3	$46,875$		

Give the 7th representative to X

Current Representation	State X	State Y	State Z
1	$281,250$	$45,000$	$76,050$
2	$93,750$		$25,350$
3	$46,875$		
4	$28,125$		

Give the 8th representative to Y

Current Representation	State X	State Y	State Z
1	$281,250$	$45,000$	$76,050$
2	$93,750$	$15,000$	$25,350$
3	$46,875$		
4	$28,125$		

Give the 9th representative to X and there is no need to recalculate because that was the last representative.

Step 6) The representatives were assigned in order to State X, Y, Z, X, X, Z, X, Y, X. The final allocation is State X has 5 representatives, State Y has 2 representatives, and State Z has 2 representatives.

Example 2

Three Universities will be sending representatives from the debate team to Washington D.C. for a special hearing. University P has 103 students on the debate team, University Q has 396 students, and University R has 247 students. Use Huntington-Hill Method to apportion 11 representatives.

Solution

The final table for Huntington-Hill numbers is below. Each University gets one representative to begin with. The number in the parentheses is the order to give out the representatives 4 through 11.

Current Representation	University P	University Q	University R
1	$5,304.5(10)$	$78,408.0(4)$	$30,504.5(5)$
2	$1,768.2$	$26,136.0(6)$	$10,168.2(8)$
3		$13,068.0(7)$	$5,084.1$
4		$7,840.8(9)$	
5		$5,227.2(11)$	

The delegates were assigned in order to University P, Q, R, Q, R, Q, Q, R, Q, P, Q. The final allocation is University P has 2 representatives, University Q has 6 representatives, and University R has 3 representatives.

CHAPTER OVERVIEW

11: Voting Systems

Thumbnail: pixabay.com/photos/survey-op...-fill-1594962/

11.1: Voting Methods

Preference Ballots

A preference ballot is a ballot in which voters rank their choices in order.

Here are four methods to determine which candidate is the winner of an election.

Plurality Method

In this method, the candidate with the most first place votes is declared the winner.

This method is sometimes mistakenly called the majority method, or "majority rules", but it is not necessary for a choice to have gained a majority of votes to win. A majority is over 50%; it is possible for a winner to have a **plurality** without having a majority.

Example 1

In our election from previous pages, we had the preference table:

	1	3	3	3
1^{st} choice	A	A	O	H
2^{nd} choice	O	H	H	A
3^{rd} choice	H	O	A	O

Solution

For the plurality method, we only care about the first choice options. Totaling them up:

Anaheim: 1+3 = 4 first-choice votes

Orlando: 3 first-choice votes

Hawaii: 3 first-choice votes

Anaheim is the winner using the plurality voting method.

Notice that Anaheim won with 4 out of 10 votes, 40% of the votes, which is a plurality of the votes, but not a majority.

Try it Now 1

Three candidates are running in an election for County Executive: Goings (G), McCarthy (M), and Bunney (B)[1]. The voting schedule is shown below. Which candidate wins under the Plurality Method?

	44	14	20	70	22	80	39
1^{st} choice	G	G	G	M	M	B	B
2^{nd} choice	M	B		G	B	M	
3^{rd} choice	B	M		B	G	G	

Note: In the third column and last column, those voters only recorded a first-place vote, so we don't know who their second and third choices would have been.

Answer

Using plurality method:

G gets $44+14+20=78$ first-choice votes

M gets $70+22=92$ first-choice votes

B gets $80+39=119$ first-choice votes

Bunney (B) wins under Plurality Method.

Plurality with Elimination is a modification of the plurality method.

Plurality with Elimination Method

In this method, the candidate with the *least* first-place votes is then eliminated from the election, and any votes for that candidate are redistributed to the voters' next choice. This continues until a candidate has a majority (over 50%).

This is similar to the idea of holding runoff elections, but since every voter's order of preference is recorded on the ballot, the runoff can be computed without requiring a second costly election.

This voting method is used in several political elections around the world, including election of members of the Australian House of Representatives, and was used for county positions in Pierce County, Washington until it was eliminated by voters in 2009. A version of IRV is used by the International Olympic Committee to select host nations.

Example 2

Consider the preference schedule below, in which a company's advertising team is voting on five different advertising slogans, called A, B, C, D, and E here for simplicity.

Initial votes

	3	4	4	6	2	1
$1^{\text{st choice}}$	B	C	B	D	B	E
2^{nd} choice	C	A	D	C	E	A
3^{rd} choice	A	D	C	A	A	D
4^{th} choice	D	B	A	E	C	B
5^{th} choice	E	E	E	B	D	C

If this was a plurality election, note that B would be the winner with 9 first-choice votes, compared to 6 for D, 4 for C, and 1 for E.

There are total of 3+4+4+6+2+1 = 20 votes. A majority would be 11 votes. No one yet has a majority, so we proceed to elimination rounds.

Solution

Round 1: We make our first elimination. Choice A has the fewest first-place votes, so we remove that choice

	3	4	4	6	2	1
1^{st} choice	B	C	B	D	B	E
2^{nd} choice	C		D	C	E	
3^{rd} choice		D	C			D
4^{th} choice	D	B		E	C	B
5^{th} choice	E	E	E	B	D	C

We then shift everyone's choices up to fill the gaps. There is still no choice with a majority, so we eliminate again.

	3	4	4	6	2	1
1^{st} choice	B	C	B	D	B	E
2^{nd} choice	C	D	D	C	E	D
3^{rd} choice	D	B	C	E	C	B
4^{th} choice	E	E	E	B	D	C

Round 2: We make our second elimination. Choice E has the fewest first-place votes, so we remove that choice, shifting everyone's options to fill the gaps.

	3	4	4	6	2	1
1^{st} choice	B	C	B	D	B	D
2^{nd} choice	C	D	D	C	C	B
3^{rd} choice	D	B	C	B	D	C

Notice that the first and fifth columns have the same preferences now, we can condense those down to one column.

	5	4	4	6	1
1^{st} choice	B	C	B	D	D
2^{nd} choice	C	D	D	C	B
3^{rd} choice	D	B	C	B	C

Now B has 9 first-choice votes, C has 4 votes, and D has 7 votes. Still no majority, so we eliminate again.

Round 3: We make our third elimination. C has the fewest votes.

	5	4	4	6	1
1^{st} choice	B	D	B	D	D
2^{nd} choice	D	B	D	B	B

Condensing this down:

	9	11
1^{st} choice	B	D
2^{nd} choice	D	B

D has now gained a majority, and is declared the winner under Plurality with Elimination.

? Try it Now 2

Consider again the election from Try it Now 1. Find the winner using Plurality with Elimination.

	44	14	20	70	22	80	39
1^{st} choice	G	G	G	M	M	B	B
2^{nd} choice	M	B		G	B	M	
3^{rd} choice	B	M		B	G	G	

Answer

G has the fewest first-choice votes, so is eliminated first. The 20 voters who did not list a second choice do not get transferred - they simply get eliminated

	136	133
1^{st} choice	M	B
2^{nd} choice	B	M

McCarthy (M) now has a majority, and is declared the winner.

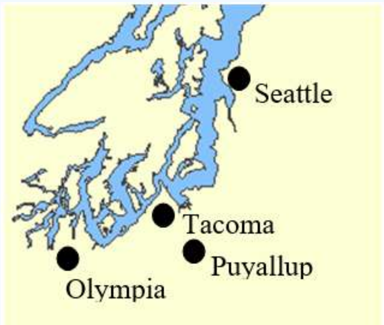
Borda Count is another voting method, named for Jean-Charles de Borda, who developed the system in 1770.

Borda Count Method

In this method, points are assigned to candidates based on their ranking; 1 point for last choice, 2 points for second-to-last choice, and so on. The point values for all ballots are totaled, and the candidate with the largest point total is the winner.

Example 3

A group of mathematicians are getting together for a conference. The members are coming from four cities: Seattle, Tacoma, Puyallup, and Olympia. Their approximate locations on a map are shown to the right.



The votes for where to hold the conference were:

	51	25	10	14
1^{st} choice	Seattle	Tacoma	Puyallup	Olympia
2^{nd} choice	Tacoma	Puyallup	Tacoma	Tacoma
3^{rd} choice	Olympia	Olympia	Olympia	Puyallup
4^{th} choice	Puyallup	Seattle	Seattle	Seattle

Solution

In each of the 51 ballots ranking Seattle first, Puyallup will be given 1 point, Olympia 2 points, Tacoma 3 points, and Seattle 4 points. Multiplying the points per vote times the number of votes allows us to calculate points awarded:

	51	25	10	14
1^{st} choice 4 points	Seattle $4 \cdot 51 = 204$	Tacoma $4 \cdot 25 = 100$	Puyallup $4 \cdot 10 = 40$	Olympia $4 \cdot 14 = 56$
2^{nd} choice 3 points	Tacoma $3 \cdot 51 = 153$	Puyallup $3 \cdot 25 = 75$	Tacoma $3 \cdot 10 = 30$	Tacoma $3 \cdot 14 = 42$
3^{rd} choice 2 points	Olympia $2 \cdot 51 = 102$	Olympia $2 \cdot 25 = 50$	Olympia $2 \cdot 10 = 20$	Puyallup $2 \cdot 14 = 28$
4^{th} choice 1 point	Puyallup $1 \cdot 51 = 51$	Seattle $1 \cdot 25 = 25$	Seattle $1 \cdot 10 = 10$	Seattle $1 \cdot 14 = 14$

Adding up the points:

- Seattle: $204 + 25 + 10 + 14 = 253$ points
- Tacoma: $153 + 100 + 30 + 42 = 325$ points
- Puyallup: $51 + 75 + 40 + 28 = 194$ points
- Olympia: $102 + 50 + 20 + 56 = 228$ points

Under the Borda Count method, Tacoma is the winner of this vote.

Try it Now 3

Consider again the election from Try it Now 1. Find the winner using Borda Count. Since we have some incomplete preference ballots, for simplicity, give every unranked candidate 1 point, the points they would normally get for last place.

	44	14	20	70	22	80	39
1^{st} choice	G	G	G	M	M	B	B
2^{nd} choice	M	B		G	B	M	
3^{rd} choice	B	M		B	G	G	

Answer

Using Borda Count:

We give 1 point for 3^{rd} place, 2 points for 2^{nd} place, and 3 points for 1^{st} place.

	44	14	20	70	22	80	39
1st choice	G 132pt	G 42pt	G 60pt	M 210pt	M 66pt	B 240pt	B 117pt
2nd choice	M 88pt	B 28pt		G 140pt	B 44pt	M 160pt	
3rd choice	B 44pt	M 14pt	M20pt B20pt	B 70pt	G 22pt	G 80pt	M39pt G39pt

G: $132+42+60+140+22+80+39=515$ pts

M: $88+14+20+210+66+160+39=597$ pts

B: $44+28+20+70+44+240+117=563$ pts

McCarthy (M) would be the winner using Borda Count.

The fourth method is the Pairwise Comparison Method.

Pairwise Comparison Method

In this method, each pair of candidates is compared, using all preferences to determine which of the two is more preferred. The more preferred candidate is awarded 1 point. If there is a tie, each candidate is awarded $\frac{1}{2}$ point. After all pairwise comparisons are made, the candidate with the most points, and hence the most pairwise wins, is declared the winner.

Example 4

Consider our vacation group example from the beginning of the chapter. Determine the winner using Pairwise Comparison Method.

	1	3	3	3
1st choice	A	A	O	H
2nd choice	O	H	H	A
3rd choice	H	O	A	O

Solution

We need to look at each pair of choices, and see which choice would win in a one-to-one comparison. For example, comparing Hawaii vs Orlando, we see that 6 voters, those shaded below in the first table below, would prefer Hawaii to Orlando. Note that Hawaii doesn't have to be the voter's first choice – we're imagining that Anaheim wasn't an option. If it helps, you can imagine removing Anaheim, as in the second table below.

	1	3	3	3
1st choice	A	A	O	H
2nd choice	O	H	H	A
3rd choice	H	O	A	O

	1	3	3	3
1st choice			O	H
2nd choice	O	H	H	
3rd choice	H	O		O

Based on this, in the comparison of Hawaii vs Orlando, Hawaii wins, and receives 1 point.

Comparing Anaheim to Orlando, the 1 voter in the first column clearly prefers Anaheim, as do the 3 voters in the second column. The 3 voters in the third column clearly prefer Orlando. The 3 voters in the last column prefer Hawaii as their first choice, but if they had to choose between Anaheim and Orlando, they'd choose Anaheim, their second choice overall. So, altogether 1+3+3=7 voters prefer Anaheim over Orlando, and 3 prefer Orlando over Anaheim. So, comparing Anaheim vs Orlando: 7 votes to 3 votes: Anaheim gets 1 point.

All together,

Hawaii vs Orlando:	6 votes to 4 votes: Hawaii gets 1 point
Anaheim vs Orlando:	7 votes to 3 votes: Anaheim gets 1 point
Hawaii vs Anaheim:	6 votes to 4 votes: Hawaii gets 1 point

Hawaii is the winner under Pairwise Comparison Method, having earned the most points.

Example 5

Consider the advertising group's vote we explored earlier. Determine the winner using Pairwise Comparison Method.

	3	4	4	6	2	1
1^{st} choice	B	C	B	D	B	E
2^{nd} choice	C	A	D	C	E	A
3^{rd} choice	A	D	C	A	A	D
4^{th} choice	D	B	A	E	C	B
5^{th} choice	E	E	E	B	D	C

Solution

With 5 candidates, there are 10 comparisons to make:

A vs B: 11 votes to 9 votes	A gets 1 point
A vs C: 3 votes to 17 votes	C gets 1 point
A vs D: 10 votes to 10 votes	A gets $\frac{1}{2}$ point, D gets $\frac{1}{2}$ point
A vs E: 17 votes to 3 votes	A gets 1 point
B vs C: 10 votes to 10 votes	B gets $\frac{1}{2}$ point, C gets $\frac{1}{2}$ point
B vs D: 9 votes to 11 votes	D gets 1 point
B vs E: 13 votes to 7 votes	B gets 1 point
C vs D: 9 votes to 11 votes	D gets 1 point
C vs E: 17 votes to 3 votes	C gets 1 point
D vs E: 17 votes to 3 votes	D gets 1 point

Totaling these up:

A gets $2\frac{1}{2}$ points

B gets $1\frac{1}{2}$ points

C gets $2\frac{1}{2}$ points

D gets $3\frac{1}{2}$ points

E gets 0 points

Using Pairwise Comparison Method, we declare D as the winner.

Try it Now 4

Consider again the election from Try it Now 1. Find the winner using Pairwise Comparison method. Since we have some incomplete preference ballots, we'll have to adjust. For example, when comparing M to B, we'll ignore the 20 votes in the third column which do not rank either candidate.

	44	14	20	70	22	80	39
1^{st} choice	G	G	G	M	M	B	B
2^{nd} choice	M	B		G	B	M	
3^{rd} choice	B	M		B	G	G	

Answer

Using Pairwise Comparison Method:

Looking back at our work from Try it Now #2, we see

G vs M: $44+14+20=78$ prefer G, $70+22+80=172$ prefer M: M preferred – 1 point

G vs B: $44+14+20+70=148$ prefer G, $22+80+39=141$ prefer B: G preferred – 1 point

M vs B: $44+70+22=136$ prefer M, $14+80+39=133$ prefer B: M preferred – 1 point

M earns 2 points; G earns 1 point. M wins under Pairwise Comparison Method.

[1] This data is loosely based on the 2008 County Executive election in Pierce County, Washington. See www.co.pierce.wa.us/xml/abtus...ec/summary.pdf

11.2: Defects in Voting Methods

Now that we have reviewed four different voting methods, how do you decide which method to use? One question to ask is which method is the fairest? Unfortunately, there is no completely fair method. What are the fairness criteria? They are guidelines that people use to help decide which voting method would be best to use under certain circumstances. They are the Majority Criterion, Condorcet Criterion, Monotonicity Criterion, and Independence of Irrelevant Alternatives Criterion.

The Majority Criterion

If a candidate receives a majority of the 1st-place votes in an election, then that candidate should be the winner of the election.

Example 1: Majority Criterion Violated

Suppose a group is planning to have a conference in one of four Arizona cities: Flagstaff, Phoenix, Tucson, or Yuma. The votes for where to hold the conference are summarized in the preference schedule shown below in Table $11.2.1$.

Table 11.2.1: Preference Schedule for Conference City

Number of voters	51	25	10	14
1st choice	Flagstaff	Phoenix	Yuma	Tucson
2nd choice	Phoenix	Yuma	Phoenix	Phoenix
3rd choice	Tucson	Tucson	Tucson	Yuma
4th choice	Yuma	Flagstaff	Flagstaff	Flagstaff

Solution

If we use the Borda Count Method to determine the winner then the number of Borda points that each candidate receives are shown in Table $11.2.2$.

Table 11.2.2: Preference Schedule for Conference City with Borda Points

Number of voters	51	25	10	14
1st choice **4 points**	Flagstaff 204	Phoenix 100	Yuma 40	Tucson 56
2nd choice **3 points**	Phoenix 153	Yuma 75	Phoenix 30	Phoenix 42
3rd choice **2 points**	Tucson 102	Tucson 50	Tucson 20	Yuma 28
4th choice **1 point**	Yuma 51	Flagstaff 25	Flagstaff 10	Flagstaff 14

The totals of all the Borda points for each city are:

Phoenix: $153 + 100 + 30 + 42 = 325$ points

Yuma: $51 + 75 + 40 + 28 = 194$ points

Tucson: $102 + 50 + 20 + 56 = 228$ points

Phoenix wins using the Borda Count Method. However, notice that Flagstaff actually has the majority of first-place votes. There are 100 voters total and 51 voters voted for Flagstaff in first place (51/100 = 51% or a majority of the first-place votes). So, Flagstaff should have won based on the Majority Criterion. This shows how the Borda Count Method can violate the Majority Criterion.

The Condorcet Criterion

If there is a candidate that in a head-to-head comparison is preferred by the voters over every other candidate, then that candidate should be the winner of the election. This candidate is known as the Condorcet candidate.

Example 2: The Condorcet Criterion Violated

Suppose you have a vacation club trying to figure out where it wants to spend next year's vacation. The choices are Hawaii (H), Anaheim (A), or Orlando (O). The preference schedule for this election is shown below in Table 11.2.3.

Table 11.2.3: Preference Schedule of Vacation Election

Number of voters	1	3	3	3
1st choice	A	A	O	H
2nd choice	O	H	H	A
3rd choice	H	O	A	O

Solution

Using the Plurality Method, A has four first-place votes, O has three first-place votes, and H has three first-place votes. So, Anaheim is the winner. However, if you use the Method of Pairwise Comparisons, A beats O (A has seven while O has three), H beats A (H has six while A has four), and H beats O (H has six while O has four). Thus, Hawaii wins all pairwise comparisons against the other candidates, and would win the election. In fact Hawaii is the Condorcet candidate. However, the Plurality Method declared Anaheim the winner, so the Plurality Method violated the Condorcet Criterion.

The Monotonicity Criterion

If candidate X is a winner of an election and, in a re-election, the only changes in the ballots are changes that favor X, then X should remain a winner of the election.

Example 3: Monotonicity Criterion Violated

Suppose you have a voting system for a mayor. The resulting preference schedule for this election is shown below in Table 11.2.4.

Table 11.2.4: Preference Schedule of Mayoral Election

Number of voters	37	22	12	29
1st choice	Adams	Brown	Brown	Carter
2nd choice	Brown	Carter	Adams	Adams
3rd choice	Carter	Adams	Carter	Brown

Solution

Using the Plurality with Elimination Method, Adams has 37 first-place votes, Brown has 34, and Carter has 29, so Carter would be eliminated. Carter's votes go to Adams, and Adams wins. Suppose that the results were announced, but then the election officials accidentally destroyed the ballots before they could be certified, so the election must be held again. Wanting to "jump on the bandwagon," 10 of the voters who had originally voted in the order Brown, Adams, Carter; change their vote to the order of Adams, Brown, Carter. No other voting changes are made. Thus, the only voting changes are in favor of Adams. The new preference schedule is shown below in Table 11.2.5.

Table 11.2.5: Preference Schedule of Mayoral Re-election

Number of voters	47	22	2	29

1st choice	Adams	Brown	Brown	Carter
2nd choice	Brown	Carter	Adams	Adams
3rd choice	Carter	Adams	Carter	Brown

Now using the Plurality with Elimination Method, Adams has 47 first-place votes, Brown has 24, and Carter has 29. This time, Brown is eliminated first instead of Carter. Two of Brown's votes go to Adams and 22 of Brown's votes go to Carter. Now, Adams has 47 + 2 = 49 votes and Carter has 29 + 22 = 51 votes. Carter wins the election. This doesn't make sense since Adams had won the election before, and the only changes that were made to the ballots were in favor of Adams. However, Adams doesn't win the re-election. The reason that this happened is that there was a difference in who was eliminated first, and that caused a difference in how the votes are re-distributed. In this example, the Plurality with Elimination Method violates the Monotonicity Criterion.

The Independence of Irrelevant Alternatives Criterion

If candidate X is a winner of an election and one (or more) of the other candidates is removed and the ballots recounted, then X should still be a winner of the election.

Example 4: The Independence of Irrelevant Alternatives Criterion Violated

A committee is trying to award a scholarship to one of four students: Anna (A), Brian (B), Carlos (C), and Dmitri (D). The votes are shown below.

Table 11.2.6: Preference Schedule for Scholarship

Number of voters	5	5	6	4
1st choice	D	A	C	B
2nd choice	A	C	B	D
3rd choice	C	B	D	A
4th choice	B	D	A	C

Solution

Using the Method of Pairwise Comparisons:

A vs B: 10 votes to 10 votes, A gets ½ point and B gets ½ point

A vs C: 14 votes to 6 votes, A gets 1 point

A vs D: 5 votes to 15 votes, D gets 1 point

B vs C: 4 votes to 16 votes, C gets 1 point

B vs D: 15 votes to 5 votes, B gets 1 point

C vs D: 11 votes to 9 votes, C gets 1 point

So A has 1½ points, B has 1 point, C has 2 points, and D has 1 point. So Carlos is awarded the scholarship.

Now suppose it turns out that Dmitri didn't qualify for the scholarship after all. Though it should make no difference, the committee decides to recount the vote. The preference schedule without Dmitri is below.

Table 11.2.7: Preference Schedule for Scholarship with Dmitri Removed

Number of voters	10	6	4
1st choice	A	C	B
2nd choice	C	B	A

3rd choice	B	A	C

Using the Method of Pairwise Comparisons:

A vs B: 10 votes to 10 votes, A gets ½ point and B gets ½ point

A vs C: 14 votes to 6 votes, A gets 1 point

B vs C: 4 votes to 16 votes, C gets 1 point

So A has 1½ points, B has ½ point, and C has 1 point. Now Anna is awarded the scholarship instead of Carlos. This is an example of The Method of Pairwise Comparisons violating the Independence of Irrelevant Alternatives Criterion.

Arrow's Impossibility Theorem

No voting system can satisfy all four fairness criteria in all cases.

In summary, every one of the fairness criteria can possibly be violated by at least one of the voting methods as shown in Table 11.2.8. However, keep in mind that this does not mean that the voting method in question will violate a criterion in every election. It is just important to know that these violations are possible.

Table 11.2.8: Summary of Violations of Fairness Criteria

	Plurality	Borda Count	Plurality with Elimination	Pairwise Comparisons
Majority Criterion	*	Violation Possible	*	*
Condorcet Criterion	Violation Possible	Violation Possible	Violation Possible	*
Monotonicity Criterion	*	*	Violation Possible	*
Independence of Irrelevant Alternatives Criterion	Violation Possible	Violation Possible	Violation Possible	Violation Possible

* *The indicated voting method does not violate the indicated criterion in any election.*

11.3: Weighted Voting

We'll begin with some basic vocabulary for weighted voting systems.

Vocabulary for Weighted Voting

Each individual or entity casting a vote is called a **player** in the election. They're often notated as $\mathbf{P_1, P_2, P_3, \ldots, P_N}$ where N is the total number of voters.

Each player is given a **weight**, which usually represents how many votes they get.

The **quota** is the minimum weight needed for the votes or weight needed for the proposal to be approved.

A weighted voting system will often be represented in a shorthand form:

$[q : w_1, w_2, w_3, \ldots, w_N]$

In this form, q is the quota, w_1 is the weight for player 1, and so on.

Example 1

In a small company, there are 4 shareholders. Mr. Smith has a 30% ownership stake in the company, Mr. Garcia has a 25% stake, Mrs. Hughes has a 25% stake, and Mrs. Lee has a 20% stake. They are trying to decide whether to open a new location. The company by-laws state that more than 50% of the ownership has to approve any decision like this. This could be represented by the weighted voting system:

$[51 : 30, 25, 25, 20]$

Solution

Here we have treated the percentage ownership as votes, so Mr. Smith gets the equivalent of 30 votes, having a 30% ownership stake. Since more than 50% is required to approve the decision, the quota is 51, the smallest whole number over 50.

In order to have a meaningful weighted voting system, it is necessary to put some limits on the quota.

Limits on Quota

The quota must be more than $\frac{1}{2}$ the total number of votes.

The quota can't be larger than the total number of votes.

Why? Consider the voting system $[q; 3, 2, 1]$

Here there are 6 total votes. If the quota was set at only 3, then player 1 could vote yes, players 2 and 3 could vote no, and both would reach quota, which doesn't lead to a decision being made. In order for only one decision to reach quota at a time, the quota must be at least half the total number of votes. If the quota was set to 7, then no group of voters could ever reach quota, and no decision can be made, so it doesn't make sense for the quota to be larger than the total number of voters.

Try it Now 1

In a committee there are four representatives from the management and three representatives from the workers' union. For a proposal to pass, four of the members must support it, including at least one member of the union. Find a voting system that can represent this situation.

Answer

If we represent the players as $M_1, M_2, M_3, M_4, U_I, U_2, U_3$ then we may be tempted to set up a system like $[4 : 1, 1, 1, 1, 1, 1, 1]$While this system would meet the first requirement that four members must support a proposal for it to pass, this does not satisfy the requirement that at least one member of the union must support it.

Consider the voting system $[10:11,3,2]$. Notice that in this system, player 1 can reach quota without the support of any other player. When this happens, we say that player 1 is a **dictator**.

Dictator

A player will be a dictator if their weight is equal to or greater than the quota. The dictator can also block any proposal from passing; the other players cannot reach quota without the dictator.

In the voting system $[8:6,3,2]$, no player is a dictator. However, in this system, the quota can only be reached if player 1 is in support of the proposal; player 2 and 3 cannot reach quota without player 1's support. In this case, player 1 is said to have **veto power**. Notice that player 1 is not a dictator, since player 1 would still need player 2 or 3's support to reach quota.

Veto Power

A player has veto power if their support is necessary for the quota to be reached. It is possible for more than one player to have veto power, or for no player to have veto power.

With the system $[10:7,6,2]$, player 3 is said to be a **dummy**, meaning they have no influence in the outcome. The only way the quota can be met is with the support of both players 1 and 2 (both of which would have veto power here); the vote of player 3 cannot affect the outcome.

Dummy

A player is a dummy if their vote is never essential for a group to reach quota.

Example 2

In the voting system $[16:7,6,3,3,2]$, are any players dictators? Do any have veto power? Are any dummies?

Solution

No player can reach quota alone, so there are no dictators.

Without player 1, the rest of the players' weights add to 14, which doesn't reach quota, so player 1 has veto power. Likewise, without player 2, the rest of the players' weights add to 15, which doesn't reach quota, so player 2 also has veto power.

Since player 1 and 2 can reach quota with either player 3 or player 4's support, neither player 3 or player 4 have veto power. However they cannot reach quota with player 5's support alone, so player 5 has no influence on the outcome and is a dummy.

Try it Now 2

In the voting system $[q:10,5,3]$, which players are dictators, have veto power, and are dummies if the quota is 10? 12? 16?

Answer

In the voting system $[q:10,5,3]$, if the quota is 10, then player 1 is a dictator since they can reach quota without the support of the other players. This makes the other two players automatically dummies.

If the quota is 12, then player 1 is necessary to reach quota, so has veto power. Since at this point either player 2 or player 3 would allow player 1 to reach quota, neither player is a dummy, so they are regular players (not dictators, no veto power, and not a dummy).

If the quota is 16, then no two players alone can reach quota, so all three players have veto power.

To better define power, we need to introduce the idea of a **coalition**. A coalition is a group of players voting the same way. In the example above, $\{P_1,P_2,P_4\}$ would represent the coalition of players 1, 2 and 4. This coalition has a combined weight of $7+6+3=16$, which meets quota, so this would be a winning coalition.

A player is said to be **critical** in a coalition if them leaving the coalition would change it from a winning coalition to a losing coalition. In the coalition $\{P_1, P_2, P_4\}$, every player is critical. In the coalition $\{P_3, P_4, P_5\}$, no player is critical, since it wasn't a winning coalition to begin with. In the coalition $\{P_1, P_2, P_3, P_4, P_5\}$, only players 1 and 2 are critical; any other player could leave the coalition and it would still meet quota.

Coalitions and Critical Players

A coalition is any group of players voting the same way.

A coalition is a **winning coalition** if the coalition has enough weight to meet quota.

A player is **critical** in a coalition if them leaving the coalition would change it from a winning coalition to a losing coalition.

Example 3

In the Scottish Parliament in 2009 there were 5 political parties: 47 representatives for the Scottish National Party, 46 for the Labour Party, 17 for the Conservative Party, 16 for the Liberal Democrats, and 2 for the Scottish Green Party. Typically all representatives from a party vote as a block, so the parliament can be treated like the weighted voting system:

$[65 : 47, 46, 17, 16, 2]$

Solution

Consider the coalition $\{P_1, P_3, P_4\}$. No two players alone could meet the quota, so all three players are critical in this coalition.

In the coalition $\{P_1, P_3, P_4, P_5\}$, any player except P_1 could leave the coalition and it would still meet quota, so only P_1 is critical in this coalition.

Notice that a player with veto power will be critical in every winning coalition, since removing their support would prevent a proposal from passing.

Likewise, a dummy will never be critical, since their support will never change a losing coalition to a winning one.

Dictators, Veto, Dummies, and Critical Players

A player is a **dictator** if the single-player coalition containing them is a winning coalition.

A player has **veto power** if they are critical in every winning coalition.

A player is a **dummy** if they are not critical in any winning coalition.

The **Banzhaf power index** was originally created in 1946 by Lionel Penrose, but was reintroduced by John Banzhaf in 1965. The power index is a numerical way of looking at power in a weighted voting situation.

Calculating Banzhaf Power Index

To calculate the Banzhaf power index:

1. List all winning coalitions
2. In each coalition, identify the players who are critical
3. Count up how many times each player is critical
4. Convert these counts to fractions or decimals by dividing by the total times any player is critical

Example 4

Find the Banzhaf power index for the voting system $[8 : 6, 3, 2]$.

Solution

We start by listing all winning coalitions. If you aren't sure how to do this, you can list all coalitions, then eliminate the non-winning coalitions. No player is a dictator, so we'll only consider two and three player coalitions.

$\{P_1, P_2\}$ Total weight: 9. Meets quota.

$\{P_1, P_3\}$ Total weight: 8. Meets quota.

$\{P_2, P_3\}$ Total weight: 5. Does not meet quota.

$\{P_1, P_2, P_3\}$ Total weight: 11. Meets quota.

Next we determine which players are critical in each winning coalition. In the winning two-player coalitions, both players are critical since no player can meet quota alone. Underlining the critical players to make it easier to count:

$\{\underline{P}_1, \underline{P}_2\}$

$\{\underline{P}_1, \underline{P}_3\}$

In the three-person coalition, either P_2 or P_3 could leave the coalition and the remaining players could still meet quota, so neither is critical. If P_1 were to leave, the remaining players could not reach quota, so P_1 is critical.

$\{\underline{P}_1, P_2, P_3\}$

Altogether, P_1 is critical 3 times, P_2 is critical 1 time, and P_3 is critical 1 time.

Converting to percents:

$P_1 = 3/5 = 60\%$

$P_2 = 1/5 = 20\%$

$P_3 = 1/5 = 20\%$

✔ Example 5

Consider the voting system $[16 : 7, 6, 3, 3, 2]$ Find the Banzhaf power index.

Solution

The winning coalitions are listed below, with the critical players underlined.

$\{\underline{P}_1, \underline{P}_2, \underline{P}_3\}$

$\{\underline{P}_1, \underline{P}_2, \underline{P}_4\}$

$\{\underline{P}_1, \underline{P}_2, P_3, P_4\}$

$\{\underline{P}_1, \underline{P}_2, \underline{P}_3, P_5\}$

$\{\underline{P}_1, \underline{P}_2, \underline{P}_4, P_5\}$

$\{\underline{P}_1, \underline{P}_2, P_3, P_4, P_5\}$

Counting up times that each player is critical:

$P_1 = 6$

$P_2 = 6$

$P_3 = 2$

$P_4 = 2$

$P_5 = 0$

Total of all: 16

Divide each player's count by 16 to convert to fractions or percents:

$P_1 = 6/16 = 3/8 = 37.5\%$

$P_2 = 6/16 = 3/8 = 37.5\%$

$P_3 = 2/16 = 1/8 = 12.5\%$

$P_4 = 2/16 = 1/8 = 12.5\%$

$P_5 = 0/16 = 0 = 0\%$

The Banzhaf power index measures a player's ability to influence the outcome of the vote. Notice that player 5 has a power index of 0, indicating that there is no coalition in which they would be critical power and could influence the outcome. This means player 5 is a dummy, as we noted earlier.

Example 6

Revisiting the Scottish Parliament, with voting system $[65 : 47, 46, 17, 16, 2]$ the winning coalitions are listed, with the critical players underlined.

Solution

$\{\underline{P_1}, \underline{P_2}\}$

$\{\underline{P_1}, \underline{P_2}, P_3\}$ $\{\underline{P_1}, \underline{P_2}, P_4\}$

$\{\underline{P_1}, \underline{P_2}, P_5\}$ $\{\underline{P_1}, \underline{P_3}, \underline{P_4}\}$

$\{\underline{P_1}, \underline{P_3}, \underline{P_5}\}$ $\{\underline{P_1}, \underline{P_4}, \underline{P_5}\}$

$\{\underline{P_2}, \underline{P_3}, \underline{P_4}\}$ $\{\underline{P_2}, \underline{P_3}, \underline{P_5}\}$

$\{P_1, P_2, P_3, P_4\}$ $\{P_1, P_2, P_3, P_5\}$

$\{\underline{P_1}, P_2, P_4, P_5\}$ $\{\underline{P_1}, P_3, P_4, P_5\}$

$\{\underline{P_2}, \underline{P_3}, P_4, P_5\}$

$\{P_1, P_2, P_3, P_4, P_5\}$

Counting up times that each player is critical:

District	Times critical	Power index
P_1 (Scottish National Party)	9	$9/27 = 33.3\%$
P_2 (Labour Party)	7	$7/27 = 25.9\%$
P_3 (Conservative Party)	5	$5/27 = 18.5\%$
P_4 (Liberal Democrats Party)	3	$3/27 = 11.1\%$
P_5 (Scottish Green Party)	3	$3/27 = 11.1\%$

Interestingly, even though the Liberal Democrats party has only one less representative than the Conservative Party, and 14 more than the Scottish Green Party, their Banzhaf power index is the same as the Scottish Green Party's. In parliamentary governments, forming coalitions is an essential part of getting results, and a party's ability to help a coalition reach quota defines its influence.

Try it Now 3

Find the Banzhaf power index for the weighted voting system $[\mathbf{36 : 20, 17, 16, 3}]$.

Answer

The voting system tells us that the quota is 36, that Player 1 has 20 votes (or equivalently, has a weight of 20), Player 2 has 17 votes, Player 3 has 16 votes, and Player 4 has 3 votes.

A coalition is any group of one or more players. What we're looking for is winning coalitions - coalitions whose combined votes (weights) add to up to the quota or more. So the coalition $\{\mathrm{P3, P4}\}$is not a winning coalition because the combined weight is $16 + 3 = 19$, which is below the quota.

So we look at each possible combination of players and identify the winning ones:

$\{\mathrm{P1}, \mathrm{P2}\}(\text{ weight } : 37)$ $\{\mathrm{P1}, \mathrm{P3}\}$ (weight: 36)
$\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}\}$ (weight: 53) $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P4}\}$ (weight: 40)
$\{\mathrm{P1}, \mathrm{P3}, \mathrm{P4}\}$ (weight: 39) $\{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P4}\}$ (weight: 56)
$\{\mathrm{P2}, \mathrm{P3}, \mathrm{P4}\}(\text{ weight: } 36)$

Example 7

Banzhaf used this index to argue that the weighted voting system used in the Nassau County Board of Supervisors in New York was unfair. The county was divided up into 6 districts, each getting voting weight proportional to the population in the district, as shown below. Calculate the power index for each district.

District	Weight
Hempstead #1	31
Hempstead #2	31
Oyster Bay	28
North Hempstead	21
Long Beach	2
Glen Cove	2

Solution

Translated into a weighted voting system, assuming a simple majority is needed for a proposal to pass:

$[58 : 31, 31, 28, 21, 2, 2]$

Listing the winning coalitions and marking critical players:

$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}\}$	$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}, \mathrm{NH}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{NH}, \mathrm{LB}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}\}$	$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}, \mathrm{LB}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{NH}, \mathrm{GC}\}$
$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}\}$	$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}, \mathrm{GC}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{LB}, \mathrm{GC}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{NH}\}$	$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}, \mathrm{NH}, \mathrm{LB}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{NH}, \mathrm{LB}, \mathrm{GC}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{LB}\}$	$\{\underline{\mathrm{H1}}, \mathrm{OB}, \mathrm{NH}, \mathrm{GC}\}$	$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{GC}\}$	$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}, \mathrm{LB}, \mathrm{GC}\}$	$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}, \mathrm{NH}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{NH}, \mathrm{LB}\}$	$\{\underline{\mathrm{H1}}, \underline{\mathrm{OB}}, \mathrm{NH}, \mathrm{LB}. \mathrm{GC}\}$	$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}, \mathrm{LB}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{NH}, \mathrm{GC}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{NH}\}$	$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}, \mathrm{GC}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{LB}, \mathrm{GC}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{LB}\}$	$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}, \mathrm{NH}, \mathrm{LB}\}$
$\{\underline{\mathrm{H1}}, \underline{\mathrm{H2}}, \mathrm{NH}, \mathrm{LB}, \mathrm{GC}\}$	$\{\underline{\mathrm{H2}}, \underline{\mathrm{OB}}, \mathrm{GC}\}$	$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}, \mathrm{NH}, \mathrm{GC}\}$
		$\{\mathrm{H1}, \mathrm{H2}, \mathrm{OB}, \mathrm{NH}, \mathrm{LB}, \mathrm{GC}\}$

There are a lot of them! Counting up how many times each player is critical,

District	Times critical	Power index
Hempstead #1	16	$16/48 = 1/3 = 33\%$
Hempstead #2	16	$16/48 = 1/3 = 33\%$
Oyster Bay	16	$16/48 = 1/3 = 33\%$
North Hempstead	0	$0/48 = 0\%$
Long Beach	0	$0/48 = 0\%$
Glen Cove	0	$0/48 = 0\%$

It turns out that the three smaller districts are dummies. Any winning coalition requires two of the larger districts.

The weighted voting system that Americans are most familiar with is the Electoral College system used to elect the President. In the Electoral College, states are given a number of votes equal to the number of their congressional representatives (house +

senate). Most states give all their electoral votes to the candidate that wins a majority in their state, turning the Electoral College into a weighted voting system, in which the states are the players.

CHAPTER OVERVIEW

12: Counting

We see probabilities almost every day in our real lives. Most times you pick up the newspaper or read the news on the internet, you encounter probability. There is a 65% chance of rain today, or a pre-election poll shows that 52% of voters approve of a ballot of the ideas students think they know about probability are incorrect. This is one area of item, are examples of probabilities. Did you ever wonder why a flush beats a full house in poker? It's because the probability of getting a flush is smaller than the probability of getting a full house. Probabilities can also be used to make business decisions, figure out insurance premiums, and determine the price of raffle tickets.

If an experiment has only three possible outcomes, does this mean that each outcome has a 1/3 chance of occurring? Many students who have not studied probability would answer yes. Unfortunately, they could be wrong. The answer depends on the experiment. Many math where their intuition is sometimes misleading. Students need to use experiments or mathematical formulas to calculate probabilities correctly.

12.1: The Fundamental Counting Principle

A tree diagram is a useful tool for visualizing systematic counting.

✓ Example 1

Let's say that a person walks into a restaurant for a three course dinner. There are four different salads, three different entrees, and two different desserts to choose from. Assuming the person wants to eat a salad, an entrée and a desert, how many different meals are possible?

Solution

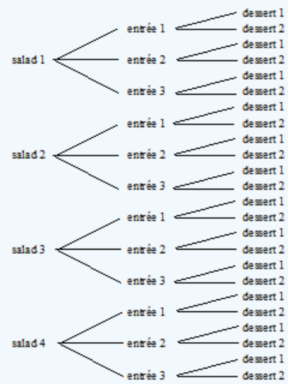


Figure 12.1.1: Tree Diagram for Three-Course Dinner

Looking at the tree diagram we can see that the total number of meals is 24. The first meal is salad 1, entrée 1, and dessert 1. The 24th meal is salad 4, entrée 3, dessert 2.

The Three Course Dinner example is easier to count the possible meals by using The Fundamental Counting Principle

✎ The Fundamental Counting Principle

If there are n_1 ways to of choosing the first item, n_2 ways of choosing the second item after the first item is chosen, n_3 ways of choosing the third item after the first two have been chosen, and so on until there are n_k ways of choosing the last item after the earlier choices, then the total number of choices overall is given by

$$n_1 \times n_2 \times n_3 \times n_4 \times n_5 \ldots \times n_k. \tag{12.1.1}$$

✓ Example 1

Let's look at the number of ways that four people can line up. We can choose any of the four people to be first. Then there are three people who can be second and two people who can be third. At this point there is only one person left to be last. Using the multiplication principle there are

$$4 \times 3 \times 2 \times 1 = 24 \text{ ways}$$

for four people to line up.

This type of calculation occurs frequently in counting problems so we have some notation to simplify the problem.

✎ Factorial

The **factorial** of n, read "n factorial" is

$$n! = n(n-1)(n-2)(n-3)\ldots(2)(1). \tag{12.1.2}$$

By this definition, $0! = 1$.

Example 2

$$\begin{aligned} 5! &= 5\times 4\times 3\times 2\times 1 \\ &= 120 \\ 8! &= 8\times 7\times 6\times 5\times 4\times 3\times 2\times 1 \\ &= 40,320 \end{aligned}$$

Factorials get very large very fast.

$$20! = 2.43\times 10^{18}$$

and

$$40! = 8.16\times 10^{47}.$$

70! is larger than most calculators can handle.

The fundamental counting principle may seem like a very simple idea but it is very powerful. Many complex counting problems can be solved using the fundamental counting principle.

Example 3

Some license plates in Arizona consist of three digits followed by three letters. How many license plates of this type are possible if:

1. There are 10 digits (0, 1, 2, 3, …, 9) and 26 letters.

$$\underbrace{(10\cdot 10\cdot 10)\cdot(26\cdot 26\cdot 26)}_{\text{digits}} = 17,576,000 \text{ license plates}$$

2. letters can be repeated but digits cannot?

$$\underbrace{(\underline{10}\cdot\underline{9}\cdot\underline{8})\cdot(26\cdot 26\cdot 26)}_{\text{letters}} = 12,654,720 \text{ license plates}$$

3. the first digit cannot be zero and both digits and letters can be repeated?

$$\underbrace{(\underline{9}\cdot\underline{10}\cdot\underline{10})}_{\text{digits}}\cdot\underbrace{(\underline{26}\cdot 26\cdot 26)}_{\text{letters}} = 15,818,400 \text{ license plates}$$

4. neither digits nor numbers can be repeated.

$$\underbrace{(\underline{10}\cdot\underline{9}\cdot\underline{8})}_{\text{digits}}\cdot\underbrace{(\underline{26}\cdot\underline{25}\cdot\underline{24})}_{\text{letters}} = 11,232,000 \text{ license plates}$$

12.2: Permutations and Combinations

Consider the following counting problems:

1. In how many ways can three runners finish a race?
2. In how many ways can a group of three people be chosen to work on a project?

What is the difference between these two problems? In the first problem the order that the runners finish the race matters. In the second problem the order in which the three people are chosen is not important, only which three people are chosen matters.

Permutation

A **permutation** is an arrangement of a set of items. The number of permutations of n items taking r at a time is given by:

$$P(n,r)=\frac{n!}{(n-r)!} \tag{12.2.1}$$

Note: Many calculators can calculate permutations directly. Look for a function that looks like ${}_nP_r$ *or* $P(n,r)$

Example 1

Let's look at a simple example to understand the formula for the number of permutations of a set of objects. Assume that 10 cars are in a race. In how many ways can three cars finish in first, second and third place? The order in which the cars finish is important. Use the multiplication principle. There are 10 possible cars to finish first. Once a car has finished first, there are nine cars to finish second. After the second car is finished, any of the eight remaining cars can finish third. 10 x 9 x 8 = 720. This is a permutation of 10 items taking three at a time.

Using the permutation formula:

$$P(10,3)=\frac{10!}{(10-3)!}=\frac{10!}{7!}=\frac{10\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}=10\cdot 9\cdot 8=720$$

Using the fundamental counting principle:

$$\underline{10}\cdot\underline{9}\cdot\underline{8}=720$$

There are 720 different ways for cars to finish in the top three places.

Example 2

The school orchestra is planning to play six pieces of music at their next concert. How many different programs are possible?

Solution

This is a permutation because they are arranging the songs in order to make the program. Using the permutation formula:

$$P(6,6)=\frac{6!}{(6-6)!}=\frac{6!}{0!}=\frac{720}{1}=720$$

Using the fundamental counting principle:

$$\underline{6}\cdot\underline{5}\cdot\underline{4}\cdot\underline{3}\cdot\underline{2}\cdot\underline{1}=720$$

There are 720 different ways of arranging the songs to make the program.

Example 3

The Volunteer Club has 18 members. An election is held to choose a president, vice-president and secretary. In how many ways can the three officers be chosen?

Solution

The order in which the officers are chosen matters so this is a permutation.

Using the permutation formula:

$$P(18,3)=\frac{18!}{(18-3)!}=\frac{18!}{15!}=18\cdot 17\cdot 16=4896$$

Note: All digits in 18! in the numerator from 15 down to one will cancel with the 15! in the denominator.

Using the fundamental principle:

$\underset{\text{Pres.}}{\underline{18}} \cdot \underset{\text{V.P.}}{\underline{17}} \cdot \underset{\text{Sec.}}{\underline{16}} = 4896$

There are 4896 different ways the three officers can be chosen.

Another notation for permutations is ${}_nP_r$. So, $P(18,3)$ can also be written as ${}_{18}P_3$. Most scientific calculators have an ${}_nP_r$ button or function.

Combinations are when the order does not even matter. We are just collecting objects together.

✓ Example 4

Choose a committee of two people from persons A, B, C, D and E. By the multiplication principle there are $5\cdot 4=20$ ways to arrange the two people.

AB AC AD AE BA BC BD BE CA CB

CD CE DA DB DC DE EA EB EC ED

Committees AB and BA are the same committee. Similarly for committees CD and DC. Every committee is counted twice.

$$\frac{20}{2}=10$$

so there are 10 possible different committees.

Now choose a committee of three people from persons A, B, C, D and E. There are $5\cdot 4\cdot 3=60$ ways to pick three people in order. Think about the committees with persons A, B and C. There are $3!=6$ of them.

ABC ACB BAC BCA CAB CBA

Each of these is counted as one of the 60 possibilities but they are the same committee. Each committee is counted six times so there are

$$\frac{60}{6}=10 \text{ different committees.}$$

In both cases we divided the number of permutations by the number of ways to rearrange the people chosen.

The number of permutations of n people taking r at a time is $P(n,r)$ and the number of ways to rearrange the people chosen is $r!$. Putting these together we get

$$\begin{aligned}\frac{n!}{\#\text{ ways to arrange r items}} &= \frac{P(n,r)}{r!}=\frac{(n-r)!}{\frac{r!}{1}}\\ &= \frac{n!}{(n-r)!}\cdot\frac{1}{r!}\\ &= \frac{n!}{(n-r)!r!}\end{aligned}$$

Combination

A **combination** is a selection of objects in which the order of selection does not matter. The number of combinations of n items taking r at a time is:

$$C(n,r)=\frac{n!}{r!(n-r)!} \tag{12.2.2}$$

Note: Many calculators can calculate combinations directly. Look for a function that looks like ${}_nC_r$ *or* $C(n,r)$.

Example 5

A student has a summer reading list of eight books. The student must read five of the books before the end of the summer. In how many ways can the student read five of the eight books?

Solution

The order of the books is not important, only which books are read. This is a combination of eight items taking five at a time.

$$C(8,5)=\frac{8!}{5!(8-5)!}=\frac{8!}{5!3!}=\frac{8\times7\times6\times5\times4\times3\times2\times1}{5\times4\times3\times2\times1\times3\times2\times1}=\frac{8\times7\times6}{3\times2\times1}=8\times7=56 \tag{12.2.3}$$

There are 56 ways to choose five of the books to read.

Example 6

A child wants to pick three pieces of Halloween candy to take in her school lunch box. If there are 13 pieces of candy to choose from, how many ways can she pick the three pieces?

Solution

This is a combination because it does not matter in what order the candy is chosen.

$$\begin{aligned}C(13,3)&=\frac{13!}{3!(13-3)!}=\frac{13!}{3!10!}\\&=\frac{13\times12\times11\times10\times9\times8\times7\times6\times5\times4\times3\times2\times1}{(3\times2\times1)(10\times9\times8\times7\times6\times5\times4\times3\times2\times1)}\\&=\frac{13\times12\times11}{3\times2\times1}\end{aligned}$$

$$=\frac{1716}{6}=286$$

There are 286 ways to choose the three pieces of candy to pack in her lunch.

Note: The difference between a combination and a permutation is whether order matters or not. If the order of the items is important, use a permutation. If the order of the items is not important, use a combination.

Now here are a couple examples where we have to figure out whether it is a permuation or a combination.

Example 7

A serial number for a particular model of bicycle consists of a letter followed by four digits and ends with two letters. Neither letters nor numbers can be repeated. How many different serial numbers are possible?

Solution

This is a permutation because the order matters.

Use the multiplication principle to solve this. There are 26 letters and 10 digits possible.

$$26\cdot10\cdot\underline{9}\cdot\underline{8}\cdot\underline{7}\cdot\underline{25}\cdot\underline{24}=78,624,000 \tag{12.2.4}$$

There are 78,624,000 different serial numbers of this form.

Example 8

A class consists of 15 men and 12 women. In how many ways can two men and two women be chosen to participate in an in-class activity?

Solution

This is a combination since the order in which the people is chosen is not important.

Choose two men:

$$C(15,2) = \frac{15!}{2!(15-2)!} = \frac{15!}{2!13!} = 105$$

Choose two women:

$$C(12,2) = \frac{12!}{2!(12-2)!} = \frac{12!}{2!10!} = 66$$

We want 2 men and 2 women so multiply these results.

$$105(66) = 6930$$

There are 6930 ways to choose two men and two women to participate.

CHAPTER OVERVIEW

13: Probability

13.1: The Basics of Probability Theory

The probability of a specified event is the chance or likelihood that it will occur. There are several ways of viewing probability. One would be **experimental** in nature, where we repeatedly conduct an experiment. Suppose we flipped a coin over and over and over again and it came up heads about half of the time; we would expect that in the future whenever we flipped the coin it would turn up heads about half of the time. When a weather reporter says “there is a 10% chance of rain tomorrow,” she is basing that on prior evidence; that out of all days with similar weather patterns, it has rained on 1 out of 10 of those days.

Another view would be **subjective** in nature, in other words an educated guess. If someone asked you the probability that the Seattle Mariners would win their next baseball game, it would be impossible to conduct an experiment where the same two teams played each other repeatedly, each time with the same starting lineup and starting pitchers, each starting at the same time of day on the same field under the precisely the same conditions. Since there are so many variables to take into account, someone familiar with baseball and with the two teams involved might make an educated guess that there is a 75% chance they will win the game; that is, *if* the same two teams were to play each other repeatedly under identical conditions, the Mariners would win about three out of every four games. But this is just a guess, with no way to verify its accuracy, and depending upon how educated the educated guesser is, a subjective probability may not be worth very much.

We will return to the experimental and subjective probabilities from time to time, but in this course we will mostly be concerned with **theoretical** probability, which is defined as follows: Suppose there is a situation with n equally likely possible outcomes and that *m* of those n outcomes correspond to a particular event; then the **probability** of that event is defined as $\frac{m}{n}$.

If you roll a die, pick a card from deck of playing cards, or randomly select a person and observe their hair color, we are executing an experiment or procedure. In probability, we look at the likelihood of different outcomes. We begin with some terminology.

Events and Outcomes

The result of an experiment is called an **outcome**.

An **event** is any particular outcome or group of outcomes.

The **sample space** is the set of all possible simple events.

Example 1

If we roll a standard 6-sided die, describe the sample space and some simple events.

The sample space is the set of all possible simple events: $\{1,2,3,4,5,6\}$

Some examples of events:

We roll a 1: outcome $\{1\}$

We roll a 5: outcome $\{5\}$

We roll a number bigger than 4: outcomes $\{5,6\}$

We roll an even number: outcomes $\{2,4,6\}$

Two dice One die

Example 2

Two dice are rolled. Write the sample space.

Solution

We assume one of the dice is red, and the other green. We have the following 36 possibilities.

Green						
Red	**1**	**2**	**3**	**4**	**5**	**6**
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)

3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

The entry (2, 5), for example, indicates that the red die shows a 2 and the green shows a 5. This is different than the entry (5, 2) which indicates that the red die shows a 5 and the green die shows a 2.

✔ Example 3

A family has three children. Write a sample space.

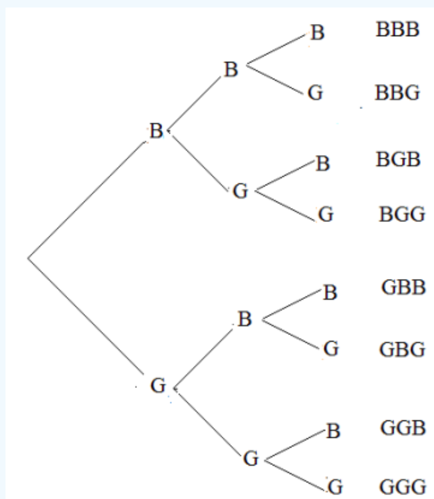
Solution

The sample space consists of eight possibilities.

{ BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG }

The possibility BGB, for example, indicates that the first born is a boy, the second born a girl, and the third a boy.

We illustrate these possibilities with a tree diagram.



✎ Basic Probability

Given that all outcomes are equally likely, we can compute the probability of an event E using this formula:

$P(E) = \frac{\text{Number of outcomes corresponding to the event E}}{\text{Total number of equally likely outcomes}}$

✔ Example 4

If we roll a 6-sided die, calculate

a. P(rolling a 1)
b. P(rolling a number bigger than 4)

Solution

Recall that the sample space is $\{1, 2, 3, 4, 5, 6\}$

a. There is one outcome corresponding to “rolling a 1”, so the probability is $\frac{1}{6}$
b. There are two outcomes bigger than a 4, so the probability is $\frac{2}{6} = \frac{1}{3}$

Probabilities are essentially fractions, and should be reduced to lowest terms.

Example 5

Let's say you have a bag with 20 cherries, 14 sweet and 6 sour. If you pick a cherry at random, what is the probability that it will be sweet?

Solution

There are 20 possible cherries that could be picked, so the number of possible outcomes is 20. Of these 20 possible outcomes, 14 are favorable (sweet), so the probability that the cherry will be sweet is $\frac{14}{20} = \frac{7}{10}$.

There is one potential complication to this example, however. It must be assumed that the probability of picking any of the cherries is the same as the probability of picking any other. This wouldn't be true if (let us imagine) the sweet cherries are smaller than the sour ones. (The sour cherries would come to hand more readily when you sampled from the bag.) Let us keep in mind, therefore, that when we assess probabilities in terms of the ratio of favorable to all potential cases, we rely heavily on the assumption of equal probability for all outcomes.

Try it Now 1

At some random moment, you look at your clock and note the minutes reading.

a. What is probability the minutes reading is 15?
b. What is the probability the minutes reading is 15 or less?

Answer

There are 60 possible readings, from 00 to 59.

a. $\frac{1}{60}$
b. $\frac{16}{60}$ (counting 00 through 15)

Deck of Playing Card

A standard deck of 52 playing cards consists of four **suits** (hearts, spades, diamonds and clubs). Spades and clubs are black while hearts and diamonds are red. Each suit contains 13 cards, each of a different **rank**: an Ace (which in many games functions as both a low card and a high card), cards numbered 2 through 10, a Jack, a Queen and a King.

Example 6

Compute the probability of randomly drawing one card from a deck and getting an Ace.

Solution

There are 52 cards in the deck and 4 Aces so

$$P(Ace) = \frac{4}{52} = \frac{1}{13} \approx 0.0769$$

We can also think of probabilities as percents: There is a 7.69% chance that a randomly selected card will be an Ace.

Notice that the smallest possible probability is 0 – if there are no outcomes that correspond with the event. The largest possible probability is 1 – if all possible outcomes correspond with the event.

Certain and Impossible Events

An **impossible event** has a probability of 0.

A **certain event** has a probability of 1.

The probability of any event must be $0 \leq P(E) \leq 1$.

In the course of this chapter, *if you compute a probability and get an answer that is negative or greater than 1, you have made a mistake and should check your work.*

13.2: Complements and Unions of Events

Complementary Events

Now let us examine the probability that an event does **not** happen. As in the previous section, consider the situation of rolling a six-sided die and first compute the probability of rolling a six: the answer is $P(\text{six}) = \frac{1}{6}$. Now consider the probability that we do *not* roll a six: there are 5 outcomes that are not a six, so the answer is $P(\text{not a six}) = \frac{5}{6}$. Notice that $P(\text{six}) + P(\text{not a six}) = \frac{1}{6} + \frac{5}{6} = \frac{6}{6} = 1$

This is not a coincidence. Consider a generic situation with n possible outcomes and an event E that corresponds to m of these outcomes. Then the remaining $n-m$ outcomes correspond to E not happening, thus $P(\text{not } E) = \frac{n-m}{n} = \frac{n}{n} - \frac{m}{n} = 1 - \frac{m}{n} = 1 - P(E)$

Complement of an Event

The **complement** of an event E is the event "E doesn't happen"

The notation $\bar{E}$ is used for the complement of event E.

We can compute the probability of the complement using $P(\bar{E}) = 1 - P(E)$

Notice also that $P(E) = 1 - P(\bar{E})$

Example 1

If you pull a random card from a deck of playing cards, what is the probability it is not a heart?

Solution

There are 13 hearts in the deck, so

$P(\text{heart}) = \frac{13}{52} = \frac{1}{4}$.

The probability of *not* drawing a heart is the complement:

$P(\text{ not heart }) = 1 - P(\text{ heart }) = 1 - \frac{1}{4} = \frac{3}{4}$

Example 2

Suppose we flipped a coin and rolled a die, and wanted to know the probability of getting a head on the coin *or* a 6 on the die.

Solution

Here, there are still 12 possible outcomes:

$\{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$

By simply counting, we can see that 7 of the outcomes have a head on the coin *or* a 6 on the die *or* both – we use *or* inclusively here (these 7 outcomes are $H1, H2, H3, H4, H5, H6, T6$), so the probability is $\frac{7}{12}$. How could we have found this from the individual probabilities?

As we would expect, $\frac{1}{2}$ of these outcomes have a head, and $\frac{1}{6}$ of these outcomes have a 6 on the die. If we add these, $\frac{1}{2} + \frac{1}{6} = \frac{6}{12} + \frac{2}{12} = \frac{8}{12}$, which is not the correct probability. Looking at the outcomes we can see why: the outcome H6 would have been counted twice, since it contains both a head and a 6; the probability of both a head *and* rolling a 6 is $\frac{1}{12}$.

If we subtract out this double count, we have the correct probability: $\frac{8}{12} - \frac{1}{12} = \frac{7}{12}$.

$P(A \text{ or } B)$

The probability of either A or B occurring (or both) is $P(A \text{ or } B)=P(A)+P(B)-P(A \text{ and } B)$

Example 3

Suppose we draw one card from a standard deck. What is the probability that we get a Queen or a King?

Solution

There are 4 Queens and 4 Kings in the deck, hence 8 outcomes corresponding to a Queen or King out of 52 possible outcomes. Thus the probability of drawing a Queen or a King is:

$P(\text{ King or Queen})=\frac{8}{52}$

Note that in this case, there are no cards that are both a Queen and a King, so $P(\text{King and Queen})=0$. Using our probability rule, we could have said:

$P(\text{ King or Queen})=P(\text{ King })+P(\text{ Queen})-P(\text{ King and Queen})=\frac{4}{52}+\frac{4}{52}-0=\frac{8}{52}$

In the last example, the events were **mutually exclusive**, so $P(A \text{ or } B)=P(A)+P(B)$.

Mutually Exclusive

Two events A and B are **mutually exclusive** if they have no outcomes in common.

Thus, $P(A \text{ or } B)=P(A)+P(B)$.

Example 4

Suppose we draw one card from a standard deck. What is the probability that we get a red card or a King?

Solution

- Half the cards are red, so $P(\text{Red})=\frac{26}{52}$
- There are four kings, so $P(\text{King})=\frac{4}{52}$
- There are two red kings, so $P(\text{Red and King})=\frac{2}{52}$

We can then calculate

$$P(\text{Red or King})=P(\text{Red})+P(\text{King})-P(\text{Red and King})=\frac{26}{52}+\frac{4}{52}-\frac{2}{52}=\frac{28}{52}=\frac{7}{13}$$

Try it Now 1

In your drawer you have 10 pairs of socks, 6 of which are white, and 7 tee shirts, 3 of which are white. If you reach in and randomly grab a pair of socks and a tee shirt, what the probability you pulled either a pair of white socks or a white tee shirt?

Answer

$P(\text{ white sock })=\frac{6}{10}$

$P(\text{ white tee })=\frac{3}{7}$

$P(\text{ white sock and white tee })=\frac{6}{10} \cdot \frac{3}{7}=\frac{9}{35}$

$P(\text{ white sock or white tee })=\frac{6}{10}+\frac{3}{7}-\frac{9}{35}=\frac{27}{35}$

13.3: Conditional Probability and Intersections of Events

Independent Events

Events A and B are **independent events** if the probability of Event B occurring is the same whether or not Event A occurs.

Example 1

Suppose we flipped a coin and rolled a die, and wanted to know the probability of getting a head on the coin and a 6 on the die.

Solution

We could list all possible outcomes: $\{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$

Notice there are $2 \cdot 6 = 12$ total outcomes. Out of these, only 1 is the desired outcome $\{H6\}$, so the probability is $\frac{1}{12}$.

Example 2

Are these events independent?

a. A fair coin is tossed two times. The two events are (1) first toss is a head and (2) second toss is a head.
b. The two events (1) "It will rain tomorrow in Houston" and (2) "It will rain tomorrow in Galveston" (a city near Houston).
c. You draw a card from a deck, then draw a second card without replacing the first.

Solution

a. The probability that a head comes up on the second toss is $\frac{1}{2}$ regardless of whether or not a head came up on the first toss, so these events are independent.
b. These events are not independent because it is more likely that it will rain in Galveston on days it rains in Houston than on days it does not.
c. The probability of the second card being red depends on whether the first card is red or not, so these events are not independent.

When two events are independent, the probability of both occurring is the product of the probabilities of the individual events.

$P(A \text{ and } B)$ for Independent Events

If events A and B are independent, then the probability of both A and B occurring is $P(A \text{ and } B) = P(A) \cdot P(B)$

If you look back at the coin and die example from earlier, you can see how the number of outcomes of the first event multiplied by the number of outcomes in the second event equals the total number of possible outcomes in the combined event.

Example 3

In your drawer you have 10 pairs of socks, 6 of which are white, and 7 tee shirts, 3 of which are white. If you randomly reach in and pull out a pair of socks and a tee shirt, what is the probability both are white?

Solution

The probability of choosing a white pair of socks is $\frac{6}{10}$.

The probability of choosing a white tee shirt is $\frac{3}{7}$.

The probability of both being white is $\frac{6}{10} \cdot \frac{3}{7} = \frac{18}{70} = \frac{9}{35}$

Try it Now 1

A card is pulled a deck of cards and noted. The card is then replaced, the deck is shuffled, and a second card is removed and noted. What is the probability that both cards are Aces?

Answer

Since the second draw is made after replacing the first card, these events are independent. The probability of an ace on each draw is $\frac{4}{52}=\frac{1}{13}$, so the probability of an Ace on both draws is $\frac{1}{13}\cdot\frac{1}{13}=\frac{1}{169}$

Example 4

The table below shows the number of survey subjects who have received and not received a speeding ticket in the last year, and the color of their car. Find the probability that a randomly chosen person:

a. Has a red car *and* got a speeding ticket
b. Has a red car *or* got a speeding ticket.

	Speeding ticket	No speeding ticket	Total
Red car	15	135	150
Not red car	45	470	515
Total	60	605	665

Solution

We can see that 15 people of the 665 surveyed had both a red car and got a speeding ticket, so the probability is $\frac{15}{665}\approx 0.0226$.

Notice that having a red car and getting a speeding ticket are not independent events, so the probability of both of them occurring is not simply the product of probabilities of each one occurring.

We could answer this question by simply adding up the numbers: 15 people with red cars and speeding tickets + 135 with red cars but no ticket + 45 with a ticket but no red car = 195 people. So the probability is $\frac{195}{665}\approx 0.2932$.

We also could have found this probability by:

$$\mathrm{P(had\ a\ red\ car)+P(got\ a\ speeding\ ticket)-P(had\ a\ red\ car\ and\ got\ a\ speeding\ ticket)}=\frac{150}{665}+\frac{60}{665}-\frac{15}{665}$$
$$=\frac{195}{665}$$

Conditional Probability

Often it is required to compute the probability of an event given that another event has occurred.

Example 5

What is the probability that two cards drawn at random from a deck of playing cards ***without replacement*** will both be aces?

Solution

It might seem that you could use the formula for the probability of two independent events and simply multiply $\frac{4}{52}\cdot\frac{4}{52}=\frac{1}{169}$. This would be incorrect, however, because the two events are not independent. If the first card drawn is an ace, then the probability that the second card is also an ace would be lower because we did not replace the first ace and therefore there would only be three aces left in the deck.

Once the first card chosen is an ace, the probability that the second card chosen is also an ace is called the **conditional probability** of drawing an ace. In this case the "condition" is that the first card is an ace. Symbolically, we write this as:

P(ace on second draw | an ace on the first draw).

The vertical bar "|" is read as "given," so the above expression is short for "The probability that an ace is drawn on the second draw given that an ace was drawn on the first draw." What is this probability? After an ace is drawn on the first draw, there are

3 aces out of 51 total cards left. This means that the conditional probability of drawing an ace after one ace has already been drawn is $\frac{3}{51}=\frac{1}{17}$.

Thus, the probability of both cards being aces is $\frac{4}{52}\cdot\frac{3}{51}=\frac{12}{2652}=\frac{1}{221}$.

Conditional Probability

The probability the event B occurs, given that event A has happened, is represented as

$P(B|A)$

This is read as “the probability of B given A”

Example 6

Find the probability that a die rolled shows a 6, given that a flipped coin shows a head.

Solution

These are two independent events, so the probability of the die rolling a 6 is $\frac{1}{6}$, regardless of the result of the coin flip.

Example 7

The table below shows the number of survey subjects who have received and not received a speeding ticket in the last year, and the color of their car. Find the probability that a randomly chosen person:

a) Has a speeding ticket *given* they have a red car

b) Has a red car *given* they have a speeding ticket

	Speeding ticket	No speeding ticket	Total
Red car	15	135	150
Not red car	45	470	515
Total	60	605	665

Solution

a) Since we know the person has a red car, we are only considering the 150 people in the first row of the table. Of those, 15 have a speeding ticket, so

P(ticket | red car) = $\frac{15}{150}=\frac{1}{10}=0.1$

b) Since we know the person has a speeding ticket, we are only considering the 60 people in the first column of the table. Of those, 15 have a red car, so

P(red car | ticket) = $\frac{15}{60}=\frac{1}{4}=0.25$.

Notice from the last example that $P(B|A)$ is **not** equal to $P(A|B)$.

These kinds of conditional probabilities are what insurance companies use to determine your insurance rates. They look at the conditional probability of you having accident, given your age, your car, your car color, your driving history, etc., and price your policy based on that likelihood.

Conditional Probability Formula

If Events A and B are not independent, then

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

Example 8

If you pull 2 cards out of a deck without replacement, what is the probability that both are spades?

Solution

The probability that the first card is a spade is $\frac{13}{52}$.

The probability that the second card is a spade, given the first was a spade, is $\frac{12}{51}$, since there is one less spade in the deck, and one less total cards.

The probability that both cards are spades is $\frac{13}{52} \cdot \frac{12}{51} = \frac{156}{2652} \approx 0.0588$

Example 9

If you draw two cards from a deck without replacement, what is the probability that you will get the Ace of Diamonds and a black card?

Solution

You can satisfy this condition by having Case A or Case B, as follows:

Case A) you can get the Ace of Diamonds first and then a black card or

Case B) you can get a black card first and then the Ace of Diamonds.

Let's calculate the probability of Case A. The probability that the first card is the Ace of Diamonds is $\frac{1}{52}$. The probability that the second card is black given that the first card is the Ace of Diamonds is $\frac{26}{51}$ because 26 of the remaining 51 cards are black. The probability is therefore $\frac{1}{52} \cdot \frac{26}{51} = \frac{1}{102}$.

Now for Case B: the probability that the first card is black is $\frac{26}{52} = \frac{1}{2}$. The probability that the second card is the Ace of Diamonds given that the first card is black is $\frac{1}{51}$. The probability of Case B is therefore $\frac{1}{2} \cdot \frac{1}{51} = \frac{1}{102}$, the same as the probability of Case 1.

Recall that the probability of A or B is $P(\text{A}) + P(\text{B}) - P(\text{A and B})$. In this problem, $P(\text{A and B}) = 0$ since the first card cannot be the Ace of Diamonds and be a black card. Therefore, the probability of Case A or Case B is $\frac{1}{102} + \frac{1}{102} = \frac{2}{102} = \frac{1}{51}$. The probability that you will get the Ace of Diamonds and a black card when drawing two cards from a deck is $\frac{1}{51}$.

Try it Now 2

In your drawer you have 10 pairs of socks, 6 of which are white. If you reach in and randomly grab two pairs of socks without replacement, what is the probability that both are white?

Answer

$\frac{6}{10} \cdot \frac{5}{9} = \frac{30}{90} = \frac{1}{3}$

Example 10

A home pregnancy test was given to women, then pregnancy was verified through blood tests. The following table shows the home pregnancy test results. Find

a) P(not pregnant | positive test result)

b) P(positive test result | not pregnant)

	Positive test	Negative test	Total
Pregnant	70	4	74
Not Pregnant	5	14	19
Total	75	18	93

Solution

a) Since we know the test result was positive, we're limited to the 75 women in the first column, of which 5 were not pregnant. P(not pregnant | positive test result) = $\frac{5}{75} \approx 0.067$.

b) Since we know the woman is not pregnant, we are limited to the 19 women in the second row, of which 5 had a positive test. P(positive test result | not pregnant) = $\frac{5}{19} \approx 0.263$.

The second result is what is usually called a false positive: A positive result when the woman is not actually pregnant.

CHAPTER OVERVIEW

14: Descriptive Statistics

14.1: Organizing and Visualizing Data

Categorical, or qualitative, data are pieces of information that allow us to classify the objects under investigation into various categories. We usually begin working with categorical data by summarizing the data into a **frequency table.**

Frequency Table

A frequency table is a table with two or three columns. One column lists the categories, and another for the frequencies with which the items in the categories occur (how many items fit into each category). The last column is the relative frequencies that give the percent of the total.

Frequency: Number of times a data value occurs in a data set.

Frequency Distribution: A listing of each data value or grouping of data values (called classes) with their frequencies.

Relative Frequency: The frequency divided by n, the size of the sample. This gives the percent of the total for each data value or class of data values.

Relative Frequency Distribution: A listing of each data value or class of data values with their relative frequencies

Example 1

An insurance company determines vehicle insurance premiums based on known risk factors. If a person is considered a higher risk, their premiums will be higher. One potential factor is the color of your car. The insurance company believes that people with some color cars are more likely to get in accidents. To research this, they examine police reports for recent total-loss collisions. The data is summarized in the frequency table below.

Color	Frequency
Blue	25
Green	52
Red	41
White	36
Black	39
Grey	23

Sometimes we need an even more intuitive way of displaying data. This is where charts and graphs come in. There are many, many ways of displaying data graphically, but we will concentrate on one very useful type of graph called a bar graph. In this section we will work with bar graphs that display categorical data; the next section will be devoted to bar graphs that display quantitative data.

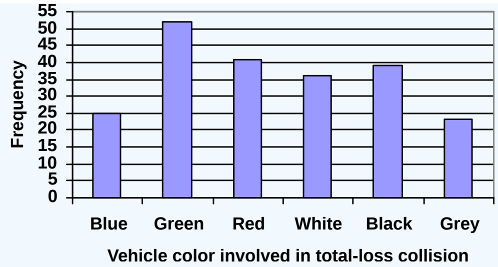
Bar Graph

A **bar graph** is a graph that displays a bar for each category with the length of each bar indicating the frequency of that category.

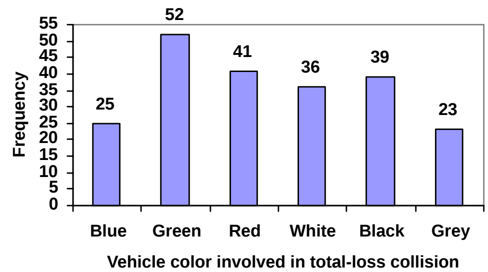
To construct a bar graph, we need to draw a vertical axis and a horizontal axis. The vertical direction will have a scale and measure the frequency of each category; the horizontal axis has no scale in this instance. The construction of a bar chart is most easily described by use of an example.

Example 2

Using our car data from above, note the highest frequency is 52, so our vertical axis needs to go from 0 to 52, but we might as well use 0 to 55, so that we can put a hash mark every 5 units:



Notice that the height of each bar is determined by the frequency of the corresponding color. The horizontal gridlines are a nice touch, but not necessary. In practice, you will find it useful to draw bar graphs using graph paper, so the gridlines will already be in place, or using technology. Instead of gridlines, we might also list the frequencies at the top of each bar, like this:



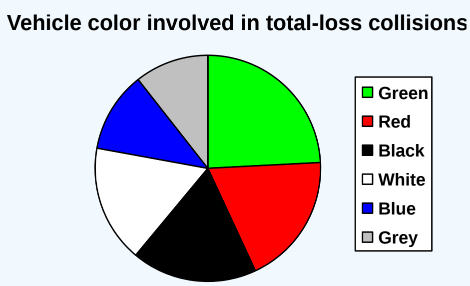
To show relative sizes, it is common to use a pie chart.

Pie Chart

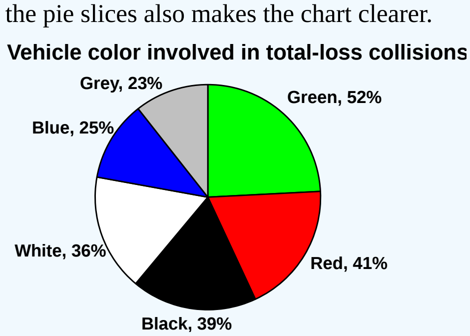
A **pie chart** is a circle with wedges cut of varying sizes marked out like slices of pie or pizza. The relative sizes of the wedges correspond to the relative frequencies of the categories.

Example 3

For our vehicle color data, a pie chart might look like this:



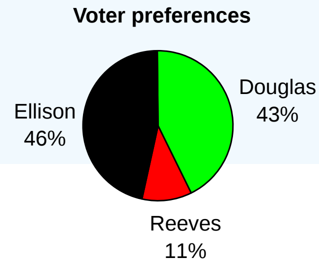
Pie charts can often benefit from including frequencies or relative frequencies (percents) in the chart next to the pie slices. Often having the category names next to the pie slices also makes the chart clearer.



✓ Example 4

The pie chart to the right shows the percentage of voters supporting each candidate running for a local senate seat.

If there are 20,000 voters in the district, the pie chart shows that about 11% of those, about 2,200 voters, support Reeves.



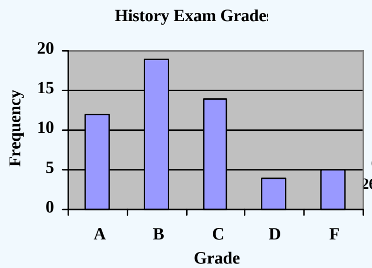
Pie charts look nice, but are harder to draw by hand than bar charts since to draw them accurately we would need to compute the angle each wedge cuts out of the circle, then measure the angle with a protractor. Computers are much better suited to drawing pie charts. Common software programs like Microsoft Word or Excel, OpenOffice.org Write or Calc, or Google Docs are able to create bar graphs, pie charts, and other graph types. There are also numerous online tools that can create graphs[2].

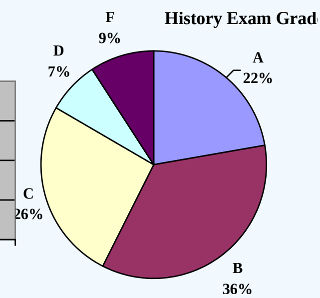
? Try it Now 1

Create a bar graph and a pie chart to illustrate the grades on a history exam below.

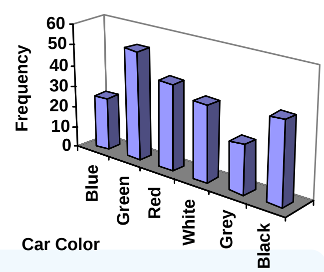
A: 12 students, B: 19 students, C: 14 students, D: 4 students, F: 5 students

Answer





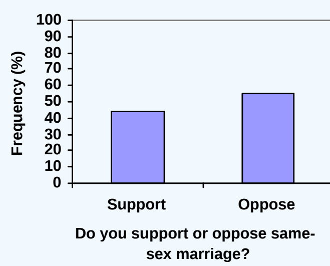
Don’t get fancy with graphs! People sometimes add features to graphs that don’t help to convey their information. For example, 3-dimensional bar charts like the one shown below are usually not as effective as their two-dimensional counterparts.

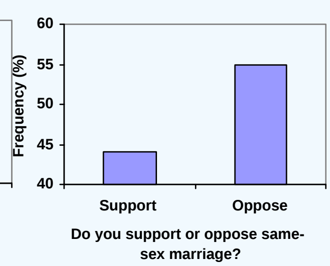


Another distortion in bar charts results from setting the baseline to a value other than zero. The baseline is the bottom of the vertical axis, representing the least number of cases that could have occurred in a category. Normally, this number should be zero.

✓ Example 5

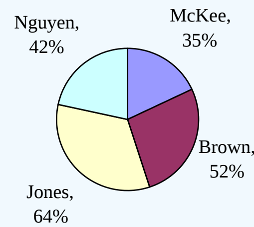
Compare the two graphs below showing support for same-sex marriage rights from a poll taken in December 2008[3]. The difference in the vertical scale on the first graph suggests a different story than the true differences in percentages; the second graph makes it look like twice as many people oppose marriage rights as support it.





Try it Now 2

A poll was taken asking people if they agreed with the positions of the 4 candidates for a county office. Does the pie chart present a good representation of this data? Explain.



Answer

While the pie chart accurately depicts the relative size of the people agreeing with each candidate, the chart is confusing, since usually percents on a pie chart represent the percentage of the pie the slice represents.

Quantitative, or numerical, data can also be summarized into frequency tables.

Example 6

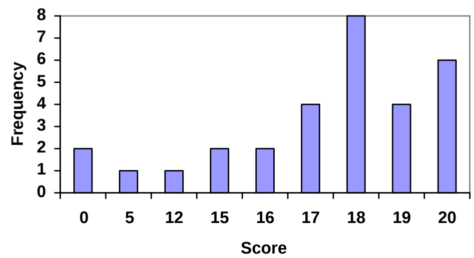
A teacher records scores on a 20-point quiz for the 30 students in his class. The scores are:

19 20 18 18 17 18 19 17 20 18 20 16 20 15 17 12 18 19 18 19 17 20 18 16 15 18 20 5 0 0

These scores could be summarized into a frequency table by grouping like values:

Score	Frequency
0	2
5	1
12	1
15	2
16	2
17	4
18	8
19	4
20	6

Using this table, it would be possible to create a standard bar chart from this summary, like we did for categorical data:



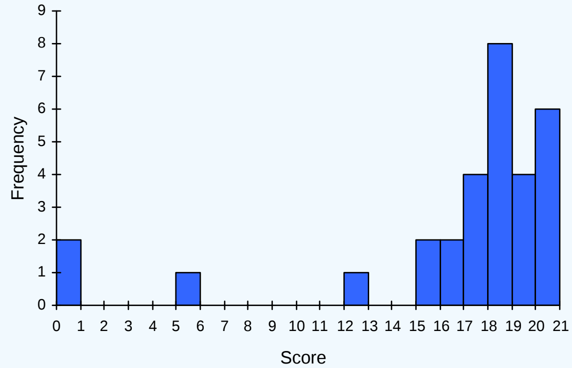
However, since the scores are numerical values, this chart doesn't really make sense; the first and second bars are five values apart, while the later bars are only one value apart. It would be more correct to treat the horizontal axis as a number line. This type of graph is called a **histogram**.

Histogram

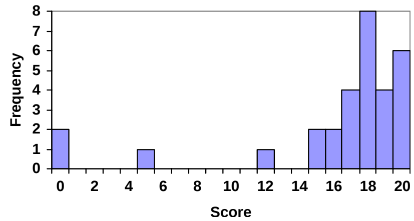
A histogram is like a bar graph, but where the horizontal axis is a number line

Example 7

For the values above, a histogram would look like:



Notice that in the histogram, a bar represents values on the horizontal axis from that on the left hand-side of the bar up to, but not including, the value on the right hand side of the bar. Some people choose to have bars start at ½ values to avoid this ambiguity.



Unfortunately, not a lot of common software packages can correctly graph a histogram. About the best you can do in Excel or Word is a bar graph with no gap between the bars and spacing added to simulate a numerical horizontal axis.

If we have a large number of widely varying data values, creating a frequency table that lists every possible value as a category would lead to an exceptionally long frequency table, and probably would not reveal any patterns. For this reason, it is common with quantitative data to group data into **class intervals**.

Class Intervals

Class intervals are groupings of the data. In general, we define class intervals so that:

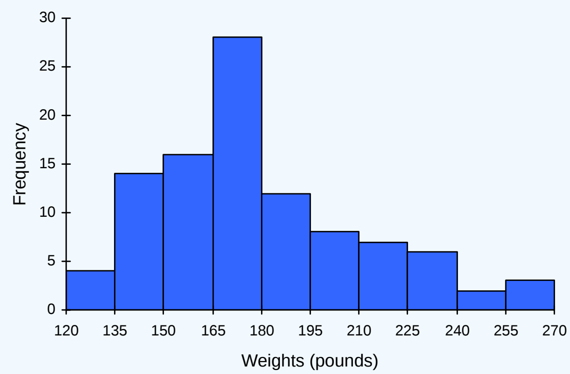
- Each interval is equal in size. For example, if the first class contains values from 120-129, the second class should include values from 130-139.
- We have somewhere between 5 and 20 classes, typically, depending upon the number of data we're working with.

Example 8

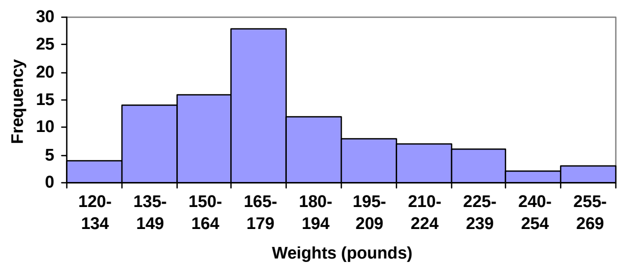
Suppose that we have collected weights from 100 male subjects as part of a nutrition study. For our weight data, we have values ranging from a low of 121 pounds to a high of 263 pounds, giving a total span of 263-121 = 142. We could create 7 intervals with a width of around 20, 14 intervals with a width of around 10, or somewhere in between. Often time we have to experiment with a few possibilities to find something that represents the data well. Let us try using an interval width of 15. We could start at 121, or at 120 since it is a nice round number.

Interval	Frequency
$120-134$	4
$135-149$	14
$150-164$	16
$165-179$	28
$180-194$	12
$195-209$	8
$210-224$	7
$225-239$	6
$240-254$	2
$255-269$	3

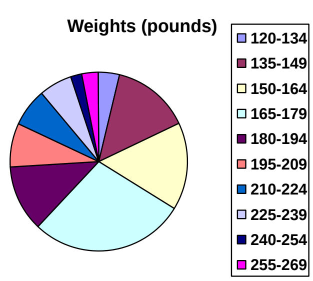
A histogram of this data would look like:



In many software packages, you can create a graph similar to a histogram by putting the class intervals as the labels on a bar chart.



Other graph types such as pie charts are possible for quantitative data. The usefulness of different graph types will vary depending upon the number of intervals and the type of data being represented. For example, a pie chart of our weight data is difficult to read because of the quantity of intervals we used.



Try it Now 3

The total cost of textbooks for the term was collected from 36 students. Create a histogram for this data.

$140 $160 $160 $165 $180 $220 $235 $240 $250 $260 $280 $285

$285 $285 $290 $300 $300 $305 $310 $310 $315 $315 $320 $320

$330 $340 $345 $350 $355 $360 $360 $380 $395 $420 $460 $460

Answer

Using a class intervals of size 55, we can group our data into six intervals:

cost interval	Frequency
$140 – 194	5
$195 – 249	3
$250 – 304	9
$305 – 359	12
$360 – 414	4
$415 – 469	3

We can use the frequency distribution to generate the histogram.

One simple graph, the *stem-and-leaf graph* or *stemplot*, comes from the field of exploratory data analysis. It is a good choice when the data sets are small. To create the plot, divide each observation of data into a stem and a leaf. The leaf consists of a final significant digit. For example, 23 has stem two and leaf three. The number 432 has stem 43 and leaf two. Likewise, the number 5,432 has stem 543 and leaf two. The decimal 9.3 has stem nine and leaf three. Write the stems in a vertical line from smallest to largest. Draw a vertical line to the right of the stems. Then write the leaves in increasing order next to their corresponding stem.

Example 9

For Susan Dean's spring pre-calculus class, scores for the first exam were as follows (smallest to largest):

33; 42; 49; 49; 53; 55; 55; 61; 63; 67; 68; 68; 69; 69; 72; 73; 74; 78; 80; 83; 88; 88; 88; 90; 92; 94; 94; 94; 94; 96; 100

Stem-and-Leaf Graph

Stem	Leaf
3	3
4	2 9 9
5	3 5 5
6	1 3 7 8 8 9 9
7	2 3 4 8
8	0 3 8 8 8
9	0 2 4 4 4 4 6
10	0

The stemplot shows that most scores fell in the 60s, 70s, 80s, and 90s. Eight out of the 31 scores or approximately 26% $\left(\frac{8}{31}\right)$ were in the 90s or 100, a fairly high number of As.

Example 10

For the Park City basketball team, scores for the last 30 games were as follows (smallest to largest):

32; 32; 33; 34; 38; 40; 42; 42; 43; 44; 46; 47; 47; 48; 48; 48; 49; 50; 50; 51; 52; 52; 52; 53; 54; 56; 57; 57; 60; 61

Construct a stem plot for the data.

Answer

Stem	Leaf
3	2 2 3 4 8
4	0 2 2 3 4 6 7 7 8 8 8 9
5	0 0 1 2 2 2 3 4 6 7 7
6	0 1

The stemplot is a quick way to graph data and gives an exact picture of the data. You want to look for an overall pattern and any outliers. An outlier is an observation of data that does not fit the rest of the data. It is sometimes called an **extreme value.** When you graph an outlier, it will appear not to fit the pattern of the graph. Some outliers are due to mistakes (for example, writing down 50 instead of 500) while others may indicate that something unusual is happening. It takes some background information to explain outliers, so we will cover them in more detail later.

Example 11

The data are the distances (in kilometers) from a home to local supermarkets. Create a stemplot using the data:

1.1; 1.5; 2.3; 2.5; 2.7; 3.2; 3.3; 3.3; 3.5; 3.8; 4.0; 4.2; 4.5; 4.5; 4.7; 4.8; 5.5; 5.6; 6.5; 6.7; 12.3

Do the data seem to have any concentration of values?

HINT: The leaves are to the right of the decimal.

Answer

The value 12.3 may be an outlier. Values appear to concentrate at three and four kilometers.

Stem	Leaf
1	1 5
2	3 5 7
3	2 3 3 5 8
4	0 2 5 5 7 8
5	5 6
6	5 7
7	
8	
9	
10	
11	
12	3

Example 12

The following data show the distances (in miles) from the homes of off-campus statistics students to the college. Create a stem plot using the data and identify any outliers:

0.5; 0.7; 1.1; 1.2; 1.2; 1.3; 1.3; 1.5; 1.5; 1.7; 1.7; 1.8; 1.9; 2.0; 2.2; 2.5; 2.6; 2.8; 2.8; 2.8; 3.5; 3.8; 4.4; 4.8; 4.9; 5.2; 5.5; 5.7; 5.8; 8.0

Answer

Stem	Leaf
0	5 7
1	1 2 2 3 3 5 5 7 7 8 9
2	0 2 5 6 8 8 8
3	5 8
4	4 8 9
5	2 5 7 8
6	
7	
8	0

The value 8.0 may be an outlier. Values appear to concentrate at one and two miles.

Example 13

A side-by-side stem-and-leaf plot allows a comparison of the two data sets in two columns. In a side-by-side stem-and-leaf plot, two sets of leaves share the same stem. The leaves are to the left and the right of the stems. Tables 14.1.1 and 14.1.2 show the ages of presidents at their inauguration and at their death. Construct a side-by-side stem-and-leaf plot using this data.

Table 14.1.1: Presidential Ages at Inauguration

President	Ageat Inauguration	President	Age	President	Age
Pierce	48	Harding	55	Obama	47
Polk	49	T. Roosevelt	42	G.H.W. Bush	64
Fillmore	50	Wilson	56	G. W. Bush	54
Tyler	51	McKinley	54	Reagan	69
Van Buren	54	B. Harrison	55	Ford	61
Washington	57	Lincoln	52	Hoover	54
Jefferson	57	Grant	46	Truman	60
Madison	57	Hayes	54	Eisenhower	62
J. Q. Adams	57	Arthur	51	L. Johnson	55
Monroe	58	Garfield	49	Kennedy	43
J. Adams	61	A. Johnson	56	F. Roosevelt	51
Jackson	61	Cleveland	47	Nixon	56

President	Ageat Inauguration	President	Age	President	Age
Taylor	64	Taft	51	Clinton	47
Buchanan	65	Coolidge	51	Trump	70
W. H. Harrison	68	Cleveland	55	Carter	52

14.1.2 *Presidential Age at Death*

President	Age	President	Age	President	Age
Washington	67	Lincoln	56	Hoover	90
J. Adams	90	A. Johnson	66	F. Roosevelt	63
Jefferson	83	Grant	63	Truman	88
Madison	85	Hayes	70	Eisenhower	78
Monroe	73	Garfield	49	Kennedy	46
J. Q. Adams	80	Arthur	56	L. Johnson	64
Jackson	78	Cleveland	71	Nixon	81
Van Buren	79	B. Harrison	67	Ford	93
W. H. Harrison	68	Cleveland	71	Reagan	93
Tyler	71	McKinley	58		
Polk	53	T. Roosevelt	60		
Taylor	65	Taft	72		
Fillmore	74	Wilson	67		
Pierce	64	Harding	57		
Buchanan	77	Coolidge	60		

Answer

Ages at Inauguration		Ages at Death
9 9 8 7 7 7 6 3 2	4	6 9
8 7 7 7 7 6 6 6 5 5 5 5 4 4 4 4 4 2 1 1 1 1 1 0	5	3 6 6 7 7 8
9 5 4 4 2 1 1 1 0	6	0 0 3 3 4 4 5 6 7 7 7 8
	7	0 0 1 1 1 4 7 8 8 9
	8	0 1 3 5 8
	9	0 0 3 3

[1] Gallup Poll. March 5-8, 2009. http://www.pollingreport.com/enviro.htm

[2] For example: http://nces.ed.gov/nceskids/createAgraph/ or http://docs.google.com

[3]CNN/Opinion Research Corporation Poll. Dec 19-21, 2008, from http://www.pollingreport.com/civil.htm

This page titled 14.1: Organizing and Visualizing Data is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman & Jeff Eldridge (The OpenTextBookStore) .

- **11.2: Presenting Categorical Data Graphically** by David Lippman & Jeff Eldridge is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.
- **5.2: Simple Interest** by Maxie Inigo, Jennifer Jameson, Kathryn Kozak, Maya Lanzetta, & Kim Sonier is licensed CC BY-SA 4.0. Original source: https://www.coconino.edu/open-source-textbooks#college-mathematics-for-everyday-life-by-inigo-jameson-kozak-lanzetta-and-sonier.
- **1.5: Graphs** by Maxie Inigo, Jennifer Jameson, Kathryn Kozak, Maya Lanzetta, & Kim Sonier is licensed CC BY-SA 4.0. Original source: https://www.coconino.edu/open-source-textbooks#college-mathematics-for-everyday-life-by-inigo-jameson-kozak-lanzetta-and-sonier.
- **11.3: Presenting Quantitative Data Graphically** by David Lippman & Jeff Eldridge is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

14.2: Measures of Central Tendency

Let's begin by trying to find the most "typical" value of a data set.

Note that we just used the word "typical" although in many cases you might think of using the word "average." We need to be careful with the word "average" as it means different things to different people in different contexts. One of the most common uses of the word "average" is what mathematicians and statisticians call the **arithmetic mean**, or just plain old **mean** for short. "Arithmetic mean" sounds rather fancy, but you have likely calculated a mean many times without realizing it; the mean is what most people think of when they use the word "average".

Mean

The **mean** of a set of data is the sum of the data values divided by the number of values.

Example 1

Marci's exam scores for her last math class were: 79, 86, 82, 94. The mean of these values would be:

Solution

$$\frac{79+86+82+94}{4}=85.25.$$

Typically we round means to one more decimal place than the original data had. In this case, we would round 85.25 to 85.3.

Example 2

The number of touchdown (TD) passes thrown by each of the 31 teams in the National Football League in the 2000 season are shown below.

37 33 33 32 29 28 28 23 22 22 22 21 21 21 20

20 19 19 18 18 18 18 16 15 14 14 14 12 12 9 6

Solution

Adding these values, we get 634 total TDs. Dividing by 31, the number of data values, we get $\frac{634}{31}=20.4516$. It would be appropriate to round this to 20.5.

It would be most correct for us to report that "The mean number of touchdown passes thrown in the NFL in the 2000 season was 20.5 passes," but it is not uncommon to see the more casual word "average" used in place of "mean.

Try it Now 1

The price of a jar of peanut butter at 5 stores was: $3.29, $3.59, $3.79, $3.75, and $3.99. Find the mean price.

Answer

The mean price is $3.68

Example 3

The one hundred families in a particular neighborhood are asked their annual household income, to the nearest $5 thousand dollars. The results are summarized in a frequency table below.

Income (thousands of dollars)	Frequency
15	6
20	8
25	11
30	17
35	19
40	20
45	12
50	7

Solution

Calculating the mean by hand could get tricky if we try to type in all 100 values:

$$\frac{15+\cdots+15+20+\cdots+20+25+\cdots+25+\cdots}{100}$$

We could calculate this more easily by noticing that adding 15 to itself six times is the same as $15 \cdot 6 = 90$. Using this simplification, we get

$$\frac{15 \cdot 6+20 \cdot 8+25 \cdot 11+30 \cdot 17+35 \cdot 19+40 \cdot 20+45 \cdot 12+50 \cdot 7}{100}=\frac{3390}{100}=33.9$$

The mean household income of our sample is 33.9 thousand dollars ($33,900).

Example 4

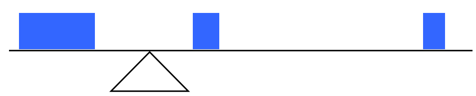
Extending off the last example, suppose a new family moves into the neighborhood example that has a household income of $5 million ($5000 thousand). Adding this to our sample, our mean is now:

Solution

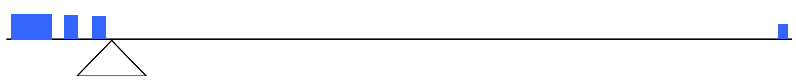
$$\frac{15 \cdot 6+20 \cdot 8+25 \cdot 11+30 \cdot 17+35 \cdot 19+40 \cdot 20+45 \cdot 12+50 \cdot 7+5000 \cdot 1}{101}=\frac{8390}{101}=83.069$$

While 83.1 thousand dollars ($83,069) is the correct mean household income, it no longer represents a “typical” value.

Imagine the data values on a see-saw or balance scale. The mean is the value that keeps the data in balance, like in the picture below.



If we graph our household data, the $5 million data value is so far out to the right that the mean has to adjust up to keep things in balance



For this reason, when working with data that have **outliers** – values far outside the primary grouping – it is common to use a different measure of center, the **median**.

Median

To find the median, begin by listing the data in order from smallest to largest.

If the number of data values, N, is odd, then the median is the middle data value. This value can be found by rounding $\frac{N}{2}$ up to the next whole number.

If the number of data values is even, there is no one middle value, so we find the mean of the two middle values (values $\frac{N}{2}$ and $\frac{N}{2}+1$)

Example 5

Returning to the football touchdown data, we would start by listing the data in order. We need to put the data in order from smallest to largest.

6 9 12 12 14 14 14 15 16 18 18 18 18 19 19 20 20 21 21 21 22 22 22 23 28 28 29 32 33 33 37

Solution

Since there are 31 data values, an odd number, the median will be the middle number, the 16th data value (\frac{31}{2} = 15.5\), round up to 16, leaving 15 values below and 15 above). The 16th data value is 20, so the median number of touchdown passes in the 2000 season was 20 passes. Notice that for this data, the median is fairly close to the mean we calculated earlier, 20.5.

Example 6

Find the median of these quiz scores: 5 10 8 6 4 8 2 5 7 7

We start by listing the data in order: 2 4 5 5 6 7 7 8 8 10

Solution

Since there are 10 data values, an even number, there is no one middle number. So we find the mean of the two middle numbers, 6 and 7, and get $\frac{6+7}{2}=6.5$.

The median quiz score was 6.5.

Try it Now 2

The price of a jar of peanut butter at 5 stores were: $3.29, $3.59, $3.79, $3.75, and $3.99. Find the median price.

Answer

First we put the data in order: $3.29, $3.59, $3.75, $3.79, $3.99. Since there are an odd number of data, the median will be the middle value, $3.75.

Example 7

Let us return now to our original household income data

Income (thousands of dollars)	Frequency
15	6
20	8
25	11
30	17
35	19
40	20
45	12
50	7

Solution

Here we have 100 data values. If we didn't already know that, we could find it by adding the frequencies. Since 100 is an even number, we need to find the mean of the middle two data values - the 50th and 51st data values. To find these, we start counting up from the bottom:

$$\begin{array}{ll} \text{There are 6 data values of \$15, so} & \text{Values 1 to 6 are \$15 thousand} \\ \text{The next 8 data values are \$20, so} & \text{Values 7 to } (6+8)=14 \text{ are \$20 thousand} \\ \text{The next 11 data values are \$25, so} & \text{Values 15 to } (14+11)=25 \text{ are \$25 thousand} \\ \text{The next 17 data values are \$30, so} & \text{Values 26 to } (25+17)=42 \text{ are \$30 thousand} \\ \text{The next 19 data values are \$35, so} & \text{Values 43 to } (42+19)=61 \text{ are \$35 thousand} \end{array} \tag{14.2.1}$$

From this we can tell that values 50 and 51 will be $35 thousand, and the mean of these two values is $35 thousand. The median income in this neighborhood is $35 thousand.

Example 8

If we add in the new neighbor with a $5 million household income, then there will be 101 data values, and the 51$^{\text{st}}$ value will be the median. As we discovered in the last example, the 51$^{\text{st}}$ value is $35 thousand. Notice that the new neighbor did not affect the median in this case. The median is not swayed as much by outliers as the mean is.

In addition to the mean and the median, there is one other common measurement of the "typical" value of a data set: the **mode**.

Mode

The **mode** is the element(s) of the data set that occurs most frequently. It is possible for a data set to have more than one mode if several categories have the same frequency, or no modes if each every category occurs only once.

The mode is fairly useless with data like weights or heights where there are a large number of possible values. The mode is most commonly used for categorical data, for which median and mean cannot be computed.

Example 9

In our vehicle color survey, we collected the data

Color	Frequency
Blue	3
Green	5
Red	4
White	3
Black	2
Grey	3

For this data, Green is the mode, since it is the data value that occurred the most frequently.

Example 10

Look back to the quiz scores example: 2 4 5 5 6 7 7 8 8 10. Find the mode of this data set.

Solution

Since 5, 7, and 8 all occur twice and that is more than any other value in the set, we say that 5, 7, and 8 are all modes.

Example 11

Marci’s exam scores for her last math class were: 79, 86, 82, 94. Find the mode of this data set.

Solution

Since none of the values are ever repeated, this data set has no mode.

Try it Now 3

Reviewers were asked to rate a product on a scale of 1 to 5. Find

a. The mean rating
b. The median rating
c. The mode rating

Rating	Frequency
1	4
2	8
3	7
4	3
5	1

Answer

a. The mean is $\frac{1\cdot 4+2\cdot 8+3\cdot 7+4\cdot 3+5\cdot 1}{23} \approx 2.5$
b. There are 23 data values, so the median will be the 12$^{\text{th}}$ data value. Ratings of 1 are the first 4 values, while a rating of 2 are the next 8 values, so the 12$^{\text{th}}$ value will be a rating of 2. The median is 2.
c. The mode is the most frequent rating. The mode rating is 2.

Quartiles

Quartiles are values that divide the data in quarters.

The first quartile (Q_1) is the value so that 25% of the data values are below it; the third quartile (Q_3) is the value so that 75% of the data values are below it. You may have guessed that the second quartile is the same as the median, since the median is the value so that 50% of the data values are below it.

This divides the data into quarters; 25% of the data is between the minimum and Q_1, 25% is between Q_1 and the median, 25% is between the median and Q_3, and 25% is between Q_3 and the maximum value

To Find Quartiles

To find the quartiles of a data set:
Step 1: Sort the data set from the smallest value to the largest value.
Step 2: Find the median (Q2). This cuts the data into two halves.
Step 3: Find the median of the lower 50% of the data values. This is the first quartile (Q1).
Step 4: Find the median of the upper 50% of the data values. This is the third quartile (Q3).

Examples should help make this clearer.

Example 12

Suppose we have measured 9 females and their heights (in inches), sorted from smallest to largest are:

59 60 62 64 66 67 69 70 72

Since the number of values is odd, the median (Q2) is the middle number 66.

To find the first quartile, we find the median of the lower half: 59 60 62 64. So Q1 = $\frac{60+62}{2}$ = 61.

To find the third quartile, we find the median of the upper half: 67 69 70 72. So Q3 = $\frac{69+70}{2}$ = 69.5

Example 13

Suppose we had measured 8 females and their heights (in inches), sorted from smallest to largest are:

59 60 62 64 66 67 69 70

Since there are an even number of data values, the median is $\frac{64+66}{2}$ = 65

To find the first quartile, we find the median of the bottom half: 59 60 62 64. So Q1 = $\frac{60+62}{2}$ = 61.

The find the third quartile, we find the median of the top half: 66 67 69 70. So Q3 = $\frac{67+69}{2}$ = 68

The 5-number summary combines the first and third quartile with the minimum, median, and maximum values.

Five Number Summary

The five number summary takes this form:

Minimum, Q_1, Median, Q_3, Maximum

Example 14

For the 9 female sample, the minimum is 59, and the maximum is 72. The 5 number summary is: 59, 61, 66, 69.5, 72.

For the 8 female sample, the minimum is 59, and the maximum is 70, so the 5 number summary would be: 59, 61, 65, 68, 70.

Try it Now 4

The total cost of textbooks for the term was collected from 36 students. Find the 5 number summary of this data.

$140 $160 $160 $165 $180 $220 $235 $240 $250 $260 $280 $285

$285 $285 $290 $300 $300 $305 $310 $310 $315 $315 $320 $320

$330 $340 $345 $350 $355 $360 $360 $380 $395 $420 $460 $460

Answer

The data is already in order, so we don’t need to sort it first.

The minimum value is $140 and the maximum is $460.

There are 36 data values so $n = 36. n/2 = 18$, which is a whole number, so the median is the mean of the 18^{th} and 19^{th} data values, $\$305$and $\$310$ The median is $\$307.50$

To find the first quartile, we find the median of the lower half:

$140 $160 $160 $165 $180 $220 $235 $240 $250 $260 $280 $285 $285 $285 $290 $300 $300 $305

Q1 = $\frac{\$250+\$260}{2}$ = $255

To find the third quartile, we find the median of the upper half:

$310 $310 $315 $315 $320 $320 $330 $340 $345 $350 $355 $360 $360 $380 $395 $420 $460 $460

Q3 = $\frac{\$345+\$350}{2}$ = $347.50

The 5 number summary of this data is: $140, $255, $307.50, $347.50, $460

Also, since we have the quartiles, we can talk about how much spread there is between the 1st and 3rd quartiles. This is known as the interquartile range.

Interquartile Range

The Interquartile Range of (IQR) = Q_3 - Q_1

For visualizing data, there is a graphical representation of a 5-number summary called a **box plot**, or box and whisker graph.

Boxplot

A **box plot** is a graphical representation of a five-number summary.

A box plot is created by first setting a scale (number line) as a guideline for the box plot. Then, draw a rectangle that spans from Q1 to Q3 above the number line. Mark the median with a vertical line through the rectangle. Next, draw dots for the minimum and maximum points to the sides of the rectangle. Finally, draw lines from the sides of the rectangle out to the dots.

Example 15

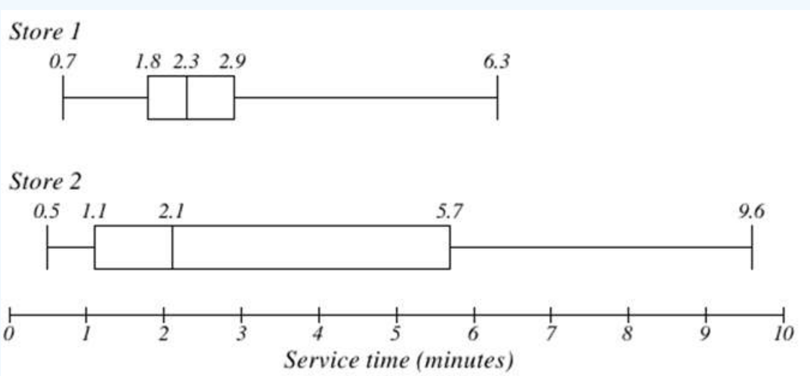
The box plot below is based on the household income data with 5 number summary:

15, 27.5, 35, 40, 50

Box plots are particularly useful for comparing data from two populations.

Example 16

The box plot of service times for two fast-food restaurants is shown below.



While store 2 had a slightly shorter median service time (2.1 minutes vs. 2.3 minutes), store 2 is less consistent, with a wider spread of the data.

At store 1, 75% of customers were served within 2.9 minutes, while at store 2, 75% of customers were served within 5.7 minutes.

Which store should you go to in a hurry? That depends upon your opinions about luck – 25% of customers at store 2 had to wait between 5.7 and 9.6 minutes.

- **11.5: Measures of Central Tendency** by David Lippman & Jeff Eldridge is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.
- **5.2: Simple Interest** by Maxie Inigo, Jennifer Jameson, Kathryn Kozak, Maya Lanzetta, & Kim Sonier is licensed CC BY-SA 4.0. Original source: https://www.coconino.edu/open-source-textbooks#college-mathematics-for-everyday-life-by-inigo-jameson-kozak-lanzetta-and-sonier.
- **11.6: Measures of Variation** by David Lippman & Jeff Eldridge is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.
- **2.3: Measures of Spread** by Maxie Inigo, Jennifer Jameson, Kathryn Kozak, Maya Lanzetta, & Kim Sonier is licensed CC BY-SA 4.0. Original source: https://www.coconino.edu/open-source-textbooks#college-mathematics-for-everyday-life-by-inigo-jameson-kozak-lanzetta-and-sonier.

14.3: Measures of Dispersion

Consider these three sets of quiz scores:

- Section A: 5 5 5 5 5 5 5 5 5 5
- Section B: 0 0 0 0 0 10 10 10 10 10
- Section C: 4 4 4 5 5 5 5 6 6 6

All three of these sets of data have a mean of 5 and median of 5, yet the sets of scores are clearly quite different. In section A, everyone had the same score; in section B half the class got no points and the other half got a perfect score, assuming this was a 10-point quiz. Section C was not as consistent as section A, but not as widely varied as section B.

In addition to the mean and median, which are measures of the "typical" or "middle" value, we also need a measure of how "spread out" or varied each data set is.

There are several ways to measure this "spread" of the data. The first is the simplest and is called the **range**.

Range

The range is the difference between the maximum value and the minimum value of the data set.

Example 1

Using the quiz scores from above,

For section A, the range is 0 since both maximum and minimum are 5 and $5-5=0$

For section B, the range is 10 since $10-0=10$

For section C, the range is 2 since $6-4=2$

In the last example, the range seems to be revealing how spread out the data is. However, suppose we add a fourth section, Section D, with scores 0 5 5 5 5 5 5 5 5 10.

This section also has a mean and median of 5. The range is 10, yet this data set is quite different than Section B. To better illuminate the differences, we'll have to turn to more sophisticated measures of variation.

Standard Deviation

The standard deviation is a measure of variation based on measuring how far each data value deviates, or is different, from the mean. A few important characteristics:

- Standard deviation is always positive. Standard deviation will be zero if all the data values are equal, and will get larger as the data spreads out.
- Standard deviation has the same units as the original data.
- Standard deviation, like the mean, can be highly influenced by outliers.

Example 2

Using the data from section D, we could compute for each data value the difference between the data value and the mean:

data value	deviation: data value - mean
0	$0-5=-5$
5	$5-5=0$
5	$5-5=0$
5	$5-5=0$
5	$5-5=0$
5	$5-5=0$
5	$5-5=0$
5	$5-5=0$
5	$5-5=0$
10	$10-5=5$

Solution

We would like to get an idea of the "average" deviation from the mean, but if we find the average of the values in the second column the negative and positive values cancel each other out (this will always happen), so to prevent this we square every value in the second column:

We then add the squared deviations up to get $25+0+0+0+0+0+0+0+0+25=50$. Ordinarily we would then divide by the number of scores, n, (in this case, 10) to find the mean of the deviations. But we only do this if the data set represents a population; if the data set represents a sample (as it almost always does), we instead divide by $n-1$ (in this case, $10-1=9$).[1]

So in our example, we would have $\frac{50}{10}=5$ if section D represents a population and $\frac{50}{9}=$ about 5.56 if section D represents a sample. These values (5 and 5.56) are called, respectively, the **population variance** and the **sample variance** for section D.

Variance can be a useful statistical concept, but note that the units of variance in this instance would be points-squared since we squared all of the deviations. What are points-squared? Good question. We would rather deal with the units we started with (points in this case), so to convert back we take the square root and get:

$$\text{population standard deviation} = \sqrt{\frac{50}{10}} = \sqrt{5} \approx 2.2$$

or

$$\text{sample standard deviation} = \sqrt{\frac{50}{9}} \approx 2.4$$

If we are unsure whether the data set is a sample or a population, we will usually assume it is a sample, and we will round answers to one more decimal place than the original data, as we have done above.

To Compute Standard Deviation

1. Find the deviation of each data from the mean. In other words, subtract the mean from the data value.
2. Square each deviation.
3. Add the squared deviations.
4. Divide by n, the number of data values, if the data represents a whole population; divide by $n-1$ if the data is from a sample.
5. Compute the square root of the result.

Example 3

Computing the standard deviation for Section B above, we first calculate that the mean is 5. Using a table can help keep track of your computations for the standard deviation:

data value	deviation: data value - mean	deviation squared
0	$0-5=-5$	$(-5)^2=25$
0	$0-5=-5$	$(-5)^2=25$
0	$0-5=-5$	$(-5)^2=25$
0	$0-5=-5$	$(-5)^2=25$
0	$0-5=-5$	$(-5)^2=25$
10	$10-5=5$	$(5)^2=25$
10	$10-5=5$	$(5)^2=25$
10	$10-5=5$	$(5)^2=25$
10	$10-5=5$	$(5)^2=25$
10	$10-5=5$	$(5)^2=25$

Assuming this data represents a population, we will add the squared deviations, divide by 10, the number of data values, and compute the square root:

$$\sqrt{\frac{25+25+25+25+25+25+25+25+25+25}{10}}=\sqrt{\frac{250}{10}}=5$$

Notice that the standard deviation of this data set is much larger than that of section D since the data in this set is more spread out.

For comparison, the standard deviations of all four sections are:

Section A: 5 5 5 5 5 5 5 5 5 5	Standard deviation: 0
Section B: 0 0 0 0 0 10 10 10 10 10	Standard deviation: 5
Section C: 4 4 4 5 5 5 5 6 6 6	Standard deviation: 0.8
Section D: 0 5 5 5 5 5 5 5 5 10	Standard deviation: 2.2

? Try it Now 1

The price of a jar of peanut butter at 5 stores were: \$3.29, \$3.59, \$3.79, \$3.75, and \$3.99. Find the standard deviation of the prices.

Answer

Earlier we found the mean of the data was \$3.682.

data value	deviation: data value - mean	deviation squared
3.29	$3.29-3.682=-0.391$	0.153664
3.59	$3.59-3.682=-0.092$	0.008464
3.79	$3.79-3.682=0.108$	0.011664
3.75	$3.75-3.682=0.068$	0.004624
3.99	$3.99-3.682=0.308$	0.094864

This data is from a sample, so we will add the squared deviations, divide by 4, the number of data values minus 1, and compute the square root:

$$\sqrt{\frac{0.153664+0.008464+0.011664+0.004624+0.094864}{4}}\approx \$0.261$$

[1] The reason we do this is highly technical, but we can see how it might be useful by considering the case of a small sample from a population that contains an outlier, which would increase the average deviation: the outlier very likely won't be included in the sample, so the mean deviation of the sample would underestimate the mean deviation of the population; thus we divide by a slightly smaller number to get a slightly bigger average deviation.

14.4: The Normal Distribution

The normal distribution is the most important of all the distributions. It is widely used and even more widely abused. Its graph is bell-shaped. You see the bell curve in almost all disciplines. Some of these include psychology, business, economics, the sciences, nursing, and, of course, mathematics. Some of your instructors may use the normal distribution to help determine your grade. Most IQ scores are normally distributed. Often real-estate prices fit a normal distribution. The normal distribution is extremely important, but it cannot be applied to everything in the real world.

Figure 14.4.1: If you ask enough people about their shoe size, you will find that your graphed data is shaped like a bell curve and can be described as normally distributed. (credit: Ömer Ünlü)

In this chapter, you will study the normal distribution, the standard normal distribution, and applications associated with them. The normal distribution has two parameters (two numerical descriptive measures), the mean (μ) and the standard deviation (σ).

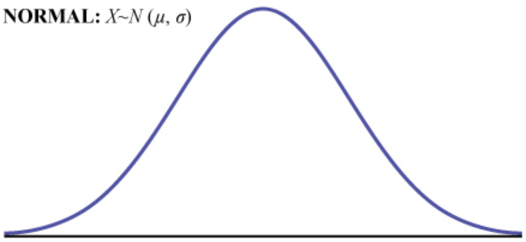


Figure 14.4.2: The standard normal distribution

The curve is symmetrical about a vertical line drawn through the mean, μ. In theory, the mean is the same as the median, because the graph is symmetric about μ. As the notation indicates, the normal distribution depends only on the mean and the standard deviation. Since the area under the curve must equal one, or 100%, a change in the standard deviation, σ, causes a change in the shape of the curve; the curve becomes fatter or skinnier depending on σ. A change in μ causes the graph to shift to the left or right. This means there are an infinite number of normal probability distributions. One of special interest is called the **standard normal distribution**.

Z-Scores

The standard normal distribution is a normal distribution of standardized values called *z-scores*. A z-score is measured in units of the standard deviation.

Z-Score

If X is a normally distributed random variable with mean μ and standard deviation σ, then the *z*-score is:

$$z = \frac{x-\mu}{\sigma}$$

The *z*-score tells you how many standard deviations the value x is above (to the right of) or below (to the left of) the mean, μ. Values of x that are larger than the mean have positive z-scores, and values of x that are smaller than the mean have negative z-scores. If x equals the mean, then x has a z-score of zero. For example, if the mean of a normal distribution is five and the standard deviation is two, the value 11 is three standard deviations above (or to the right of) the mean. The calculation is as follows:

$$\begin{aligned} x &= \mu + (z)(\sigma) \\ &= 5 + (3)(2) = 11 \end{aligned}$$

The *z*-score is three.

Since the mean for the standard normal distribution is zero and the standard deviation is one, then this transformation produces the standard normal distribution.

Example 1

Suppose that x is a normally distributed random variable with mean $\mu = 5$ and standard deviation $\sigma = 6$. Suppose $x = 17$. Then the z-score is $z = \dfrac{x-\mu}{\sigma} = \dfrac{17-5}{6} = 2$

This means that $x = 17$ is **two** standard deviations (2σ) above or to the right of the mean $\mu = 5$. The standard deviation is $\sigma = 6$.

Now suppose $x = 1$. Then the z-score is $z = \dfrac{x-\mu}{\sigma} = \dfrac{1-5}{6} = -0.67$ (rounded to two decimal places)

This means that $x = 1$ is 0.67 standard deviations (-0.67σ) below or to the left of the mean $\mu = 5$.

Summarizing, when z is positive, x is above or to the right of μ and when z is negative, x is to the left of or below μ. Or, when z is positive, x is greater than μ, and when z is negative x is less than μ.

Try it Now 1

What is the z-score of x, when $x = 1$ and mean $\mu = 12$ and standard deviation $\sigma = 3$?

Answer

$$z = \frac{1-12}{3} \approx -3.67$$

Example 2

Some doctors believe that a person can lose five pounds, on the average $\mu = 5$, in a month by reducing his or her fat intake and by exercising consistently. Suppose weight loss has a normal distribution. Let $X =$ the amount of weight lost(in pounds) by a person in a month. Use a standard deviation of two pounds $\sigma = 2$. Fill in the blanks.

a. Suppose a person **lost** ten pounds in a month. The z-score when $x = 10$ pounds is $x = 2.5$ (verify). This z-score tells you that $x = 10$ is ________ standard deviations to the ________ (right or left) of the mean _____ (What is the mean?).
b. Suppose a person **gained** three pounds (a negative weight loss). Then $z =$ __________. This z-score tells you that $x = -3$ is ________ standard deviations to the __________ (right or left) of the mean.

Answers

a. This z-score tells you that $x = 10$ is 2.5 standard deviations to the right of the mean five.

b. Then $z = \dfrac{-3-5}{2} = -4$. This z-score tells you that $x = -3$ is 4 standard deviations to the left of the mean five.

Try it Now 2

Fill in the blanks.

Jerome averages 16 points a game, $\mu = 16$, with a standard deviation of four points, $\sigma = 4$. Suppose Jerome scores ten points in a game. The z–score when $x = 10$ is -1.5. This score tells you that $x = 10$ is _____ standard deviations to the ______(right or left) of the mean______(What is the mean?).

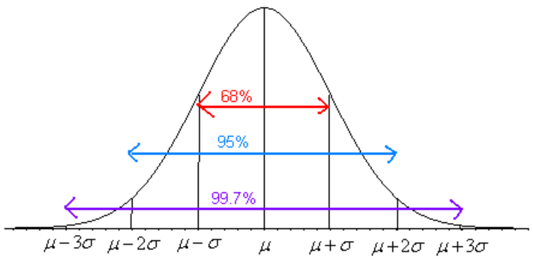
Answer

1.5, left, 16

The Empirical Rule (68-95-99.7 Rule)

If X is a random variable and has a normal distribution with mean μ and standard deviation σ, then the *Empirical Rule* says the following:

- About 68% of the x values lie between –1σ and +1σ of the mean μ (within one standard deviation of the mean).
- About 95% of the x values lie between –2σ and +2σ of the mean μ (within two standard deviations of the mean).
- About 99.7% of the x values lie between –3σ and +3σ of the mean μ (within three standard deviations of the mean). Notice that almost all the x values lie within three standard deviations of the mean.
- The z-scores for +1σ and –1σ are +1 and –1, respectively.
- The z-scores for +2σ and –2σ are +2 and –2, respectively.
- The z-scores for +3σ and –3σ are +3 and –3 respectively.



Example 3

Suppose *x* has a normal distribution with mean 50 and standard deviation 6.

- About 68% of the *x* values lie within one standard deviation of the mean. Therefore, about 68% of the *x* values lie between –1σ = (–1)(6) = –6 and 1σ = (1)(6) = 6 of the mean 50. The values 50 – 6 = 44 and 50 + 6 = 56 are within one standard deviation from the mean 50. The z-scores are –1 and +1 for 44 and 56, respectively.
- About 95% of the x values lie within two standard deviations of the mean. Therefore, about 95% of the x values lie between –2σ = (–2)(6) = –12 and 2σ = (2)(6) = 12. The values 50 – 12 = 38 and 50 + 12 = 62 are within two standard deviations from the mean 50. The z-scores are –2 and +2 for 38 and 62, respectively.
- About 99.7% of the x values lie within three standard deviations of the mean. Therefore, about 99.7% of the x values lie between –3σ = (–3)(6) = –18 and 3σ = (3)(6) = 18 from the mean 50. The values 50 – 18 = 32 and 50 + 18 = 68 are within three standard deviations of the mean 50. The z-scores are –3 and +3 for 32 and 68, respectively.

Try it Now 3

Suppose X has a normal distribution with mean 25 and standard deviation five. Between what values of x do 68% of the values lie?

Answer

between 20 and 30.

✓ Example 4

From 1984 to 1985, the mean height of 15 to 18-year-old males from Chile was 172.36 cm, and the standard deviation was 6.34 cm. Let $Y =$ the height of 15 to 18-year-old males in 1984 to 1985.

a. About 68% of the y values lie between what two values? These values are ________________. The z-scores are ________________, respectively.
b. About 95% of the y values lie between what two values? These values are ________________. The z-scores are ________________ respectively.
c. About 99.7% of the y values lie between what two values? These values are ________________. The z-scores are ________________, respectively.

Answer

a. About 68% of the values lie between 166.02 and 178.7. The z-scores are –1 and 1.
b. About 95% of the values lie between 159.68 and 185.04. The z-scores are –2 and 2.
c. About 99.7% of the values lie between 153.34 and 191.38. The z-scores are –3 and 3.

? Try it Now 4

The scores on a college entrance exam have an approximate normal distribution with mean, $\mu = 52$ points and a standard deviation, $\sigma = 11$ points.

a. About 68% of the y values lie between what two values? These values are ________________. The z-scores are ________________, respectively.
b. About 95% of the y values lie between what two values? These values are ________________. The z-scores are ________________, respectively.
c. About 99.7% of the y values lie between what two values? These values are ________________. The z-scores are ________________, respectively.

Answer a

About 68% of the values lie between the values 41 and 63. The z-scores are –1 and 1, respectively.

Answer b

About 95% of the values lie between the values 30 and 74. The z-scores are –2 and 2, respectively.

Answer c

About 99.7% of the values lie between the values 19 and 85. The z-scores are –3 and 3, respectively.

Index

Glossary

Under | Construction

Detailed Licensing

Overview

Title: Math for Liberal Arts

Webpages: 85

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-SA 4.0: 64.7% (55 pages)
- Undeclared: 22.4% (19 pages)
- CC BY-NC-SA 4.0: 7.1% (6 pages)
- CC BY-SA 3.0: 3.5% (3 pages)
- CC BY-NC-SA 3.0: 1.2% (1 page)
- CC BY 4.0: 1.2% (1 page)

By Page

- Math for Liberal Arts - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Problem Solving - *CC BY-SA 4.0*
    - 1.1: Introduction - *Undeclared*
    - 1.2: Percents - *CC BY-SA 4.0*
    - 1.3: Proportions and Rates - *CC BY-SA 4.0*
    - 1.4: Geometry - *CC BY-SA 4.0*
    - 1.5: Problem Solving and Estimating - *CC BY-SA 4.0*
    - 1.6: Exercises - *CC BY-SA 4.0*
    - 1.7: Extension - Taxes - *CC BY-SA 4.0*
    - 1.8: Income Taxation - *Undeclared*
  - 2: Set Theory - *CC BY-SA 4.0*
    - 2.1: Language of Sets - *CC BY-SA 4.0*
    - 2.2: Comparing Sets - *CC BY-SA 4.0*
    - 2.3: Set Operations - *CC BY-SA 4.0*
    - 2.4: Survey Problems - *CC BY-SA 4.0*
  - 3: Logic - *CC BY-SA 4.0*
    - 3.1: Statements, Connectives, and Quantifiers - *CC BY-SA 4.0*
    - 3.2: Truth Tables- Conjunction (and), Disjunction (or), Negation (not) - *CC BY-SA 4.0*
    - 3.3: Truth Tables- Conditional, Biconditional - *CC BY-SA 4.0*
    - 3.4: Verifying Arguments - *Undeclared*
  - 4: Graph Theory - *CC BY-SA 4.0*
    - 4.1: Graphs - *CC BY-SA 4.0*
    - 4.2: Hamiltonian Circuits and the Traveling Salesman Problem - *CC BY-SA 4.0*
    - 4.3: Shortest Path - *CC BY-SA 4.0*
    - 4.4: Spanning Trees - *CC BY-SA 4.0*
  - 5: Numeration Systems - *CC BY-SA 4.0*
    - 5.1: The Evolution of Numeration Systems - *CC BY-SA 4.0*
    - 5.2: Place Value Systems - *CC BY-SA 4.0*
    - 5.3: Calculating in Other Bases - *CC BY-SA 4.0*
    - 5.4: Modular Systems - *CC BY-NC-SA 4.0*
  - 6: Cryptography - *CC BY-SA 4.0*
    - 6.1: Cryptography - *CC BY-SA 4.0*
    - 6.2: Substitution Ciphers - *CC BY-SA 4.0*
    - 6.3: Transposition Ciphers - *CC BY-SA 4.0*
    - 6.4: Advanced shared symmetric-key methods - *CC BY-SA 4.0*
    - 6.5: Public Key Cryptography - *CC BY-SA 4.0*
    - 6.6: Other Secret Keeping Methods - *CC BY-SA 4.0*
    - 6.7: The One-Way Function - *Undeclared*
    - 6.8: The key exchange - *Undeclared*
  - 7: Algebraic Models - *CC BY-SA 4.0*
    - 7.1: Linear Equations - *Undeclared*
    - 7.2: Modeling with Linear Equations - *CC BY-SA 4.0*
    - 7.3: Modeling with Quadratic Equations - *CC BY-SA 4.0*
    - 7.4: Exponential Growth - *CC BY-SA 4.0*
  - 8: Consumer Mathematics - *CC BY-SA 4.0*
    - 8.1: Percent - *Undeclared*
    - 8.2: Simple and Compound Interest - *CC BY-SA 4.0*
    - 8.3: Credit Cards - *Undeclared*
    - 8.4: Annuities - *CC BY-SA 4.0*
    - 8.5: Amortized Loans - *CC BY-SA 4.0*
  - 9: Geometry - *CC BY-NC-SA 4.0*
    - 9.1: Lines and Angles - *CC BY-NC-SA 4.0*