# Math Excursions
# for
# Liberal Arts

Ben Weng

Pangyen (Ben) Weng, PhD
Minneapolis, Minnesota, USA
Email: drweng.net@gmail.com
URL: http://drweng.net

# *Contents*

# *Foreword*

You don't have to be an athlete to enjoy an exciting ballgame, or a musician to enjoy a good song, or an artist to enjoy a beautiful painting. You also don't need to be a mathematician to enjoy mathematics.

This book is an exploration into interesting topics and intriguing discussions of mathematics. You won't find algebra drills in it, but an abundant supply of mathematical ideas, thought-provoking conversations, and examples and applications of math in the modern world.

Bon voyage.

## *To Fellow Teachers*

The Math for Liberal Arts course is college education's best (or last?) opportunity to inspire the non STEM students with math. This book is a collection of topics and activities that work well when I teach this course. It covers many topics normally not taught in a conventional math course, and it uses examples and discussions on real world issues such as COVID-19, social inequity, negative political campaigning, educational disparity, professional sports, and much more. My students love them, and so will yours.

Not all of the weekly videos align with the chapter content: I sometimes choose to give the students a break by assigning interesting videos on other topics. Except for logic and statistics, the chapters are independent from each other and do not need to be taught in order. Two chapters are lighter than others: Chapter One is intended for the opening week, and Chapter Ten can be used for a short week such as Thanksgiving or the week before spring break.

This book is written for 8-week or 12-week classes: I teach seven chapters plus one project in 8 weeks, or ten chapters plus two projects in 12 weeks. The students are expected to read all the materials independently: The content is friendly enough, and can be enforced with flipped teaching methods. The class time is dedicated to discussions, and for online classes I use discussion forums. I assign grades based on the project(s), the discussions and the weekly multiple-choice quizzes that focus on comprehension instead of algebra skills. Feel free to email me about a sample project or a sample quiz at DrWeng.Net@gmail.com.

I wish you and your students a fantastic journey into the mathematics for liberal arts.

## *Acknowledgement*

This book wouldn't be possible without Dr. Cindy Kaus of Metropolitan State University in St. Paul, Minnesota. Cindy is a great math educator and a dear mentor and colleague to me. Cindy helped my team build the Math for Liberal Arts course by sharing her curriculum and pedagogy. Many topics in this book are inspired by her teaching philosophy, and the three chapters in statistics are derived from her class notes.

I am grateful to my supervisor Dr. Gail O'Kane for supporting my team's effort to reform the general education math at our college. The Math for Liberal Arts course is a crucial part of the Minneapolis Math Pathways, and I am lucky to participate as its dean and one of its teachers.

I also thank Becky Groseth of Anoka-Ramsey Community College for providing me with valuable suggestions and for catching numerous errors and typos in the 2021 edition (which you should NEVER buy).

Last but not the least, I thank my beautiful wife Emily and our three kids Euclid, Belle and Leo. They are the only people in my universe who'd frankly tell me when my stuff is boring. (In Leo's case, he screams and asks to go out for ice cream.) If this book is interesting to the readers, it is because of their push for me to work on my storytelling and message-crafting skills.

Now, on to the mathematics.

*Chapter 1*

# *Exploring the World of Numbers*

## 1.1 Video: The Story of One

Watch this documentary on YouTube: The Story of One, BBC Documentary 2014

*Discussions*

1. **What are numbers?** An alien just landed on the Earth and asked you about "numbers". Explain numbers in less than 50 words.

2. **Numbers in ancient civilizations.** How did Sumerians and Egyptians use their numbers? Base your discussion on the video and other sources of info you find. You must refer to your sources, like "This is from the video 15:00–18:00" or "This is on this website I found: …".

3. **Mathematicians.** Choose one of the mathematicians mentioned in this video. Find out more about this person, their discoveries and their influences on the human civilization. Base your discussion on the video and other viable sources of information. Make sure to refer to your sources.

4. **Indian mathematics.** What are the great discoveries of Indian mathematics? What influence do these discoveries have on the human civilization?

5. **The binary numbers.** Who first introduced the binary number system and why?

6. **The first computer.** According to the video, what is the first computer and what did it do? How does the binary number system play a role in building computers?

## 1.2 Computers and Binary Codes

### *The Product Rule of Counting*

Assume that something has two parts. If there are $m$ options for one part and $n$ options for the other part, then there are $m \times n$ options for the entire thing. For example, if there are 5 flavors of ice cream and 3 sizes of cones, then there are $5 \times 3 = 15$ different ways to make an ice cream cone with one scoop on a cone.

If there are more than two parts, we simply multiply the numbers of options for each part. For example, assume a meal includes a burger, a side and a beverage. If a restaurant serves 5 kinds of burgers, 3 kinds of sides and 4 kinds of beverages, then the restaurant can serve $5 \times 4 \times 3 = 60$ different kinds of meals.

### *Bits and Bytes in Computers*

Computers use electric circuits to store data and to perform calculations. The basic circuit is a single switch called a *bit*. A bit is either off or on, which are denoted by 0 (off) and 1 (on). The *bytes* are the basic data units in computers. A byte consists of 8 bits and looks like □□□□□□□□ with 0 or 1 in each box. For example, 00011001 and 11110010 are both data that can be saved in a byte. How many different data can be saved in a byte? Since a byte has 8 bits and each bit has two options (0 or 1), the number of options to write in a byte is $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 = 256$, by the product rule of counting.

One way to use the 256 different data of a byte is to code them for alphabets, numerals and symbols (aka *characters*), and then we can convert articles or even books to binary codes. A common practice is via the extended *American Standard Code for Information Interchange*, or the (extended) ASCII codes. You can find a table for these codes at https://www.ascii-code.com. On the table, for example, you will find that 01001001 is the code for the uppercase letter I. Can you decode this following message?

| | |
|---|---|
| 01001001 | 00100000 |
| 01101100 | 01101111 |
| 01110110 | 01100101 |
| 00100000 | 01111001 |
| 01101111 | 01110101 |

### *What Color Is Your HDTV?*

The 256 possibilities of a byte are also used for digital graphics. In digital images, any color is a combination of three colors: red, green, and blue.

An 8-bit color scheme uses 0 (none) to 255 (maximum color) for the level of each of the three colors. Since any color is made of the three colors and each color has 256 options, the number of colors that can be created is $256 \times 256 \times 256 = 16,777,216$, or over 16 million colors. These are the so-called *8-bit colors*.

What if we wish to have even more colors and add more *depth* to the colors? Since the 2010s, more and more high-definition televisions (HDTVs) have been using the 10-bit and 12-bit colors. For the 10-bit colors, the three colors (red, green and blue) are recorded as 10-bit data, giving them $2^{10} = 1024$ different levels. Thus, there are 1024 options for each of the three colors, and the number of options for the overall color is $1024 \times 1024 \times 1024 = 1,073,741,824$, or over 1 billion colors.

Can you tell the difference between 8-bit colors and 10-bit colors? Visit a local electronic store and find out!

### *Discussions*

1. **Minnesota license plates.** A format of car license plates in Minnesota is *three uppercase letters followed by three numerals*, like ABC-999. The letters and numerals may repeat. How many different license plates can be formed in this format?

2. **Decoding.** Finish decoding the message in the lesson.

3. **12-bit colors.** How many 12-bit colors are there? You may use a calculator but explain your thoughts.

## *1.3* $1 + 1 = 2$, *Right?*

### *Does Math Dictate the Reality?*

Mathematics is an abstract form to describe the reality. It states the general principles, and it deduces what would happen, based on logic and these general principles. However, it is *the principles* that dictate what happens, not the mathematical formulas.

While most laws in the world are validated by human institutions, mathematical laws are valid because they describe *how things actually work*. For example, there is a law in algebra called the *distributive law*: If $a$, $b$ and $c$ are three numbers, then $a \times (b + c) = a \times b + a \times c$. Let's demonstrate how the formula works by considering a special case $5 \times (8 + 2) = 5 \times 8 + 5 \times 2$. Assume a food truck sells lunch meals that include one burrito (\$8) and one beverage (\$2). So five meals ($5 \times \$10$) is equal to five burritos ($5 \times \$8$) and five beverages ($5 \times \$2$). A general proof can be constructed using the same idea.

What if we want to propose a similar formula for a mathematical law: $\frac{a}{b+c} = \frac{a}{b} + \frac{a}{c}$? Unfortunately, this won't be valid because things simply don't work that way. For example, $\frac{6}{1+2} = 2$, which is different from $\frac{6}{1} + \frac{6}{2} = 6 + 3 = 9$. No matter how good this formula looks, it can't be a mathematical law.

*The Assumptions Matter*

While mathematics describes things in general, we should avoid losing the context or *the assumptions* for the math to apply, even for something as basic as $1 + 1 = 2$. With $1 + 1 = 2$, we mean such things as "1 cookie and 1 cookie make 2 cookies", "1 gallon of water and 1 gallon of water make 2 gallons of water", etc. The underlying assumption is that we are combining the same kind of stuff. Without this assumption, one may tweak the meaning of $1 + 1$ with totally different stories. For example, if a wolf and a chicken are put in the same cage, the result would be a well-fed wolf and no chicken, so $1 + 1$ is definitely not 2.

Another example is the commutative law of addition, which says that *when adding multiple numbers, you can alter the order of these numbers and still get the same result.* This is correct in the abstract, but not always so in the context of financial transactions. Consider the following two scenarios.

1. $500 + (-600) + 900$: You have \$500 in cash, a \$600 rent due, and later receive a \$900 payment for a job.

2. $500 + 900 + (-600)$: You have \$500 in cash, receive a \$900 payment for a job, and then have a \$600 rent due.

In the first scenario, you won't be able to pay the rent on time, which might result in a penalty or a lowered credit score. On the other hand, you are perfectly fine in the second scenario. This shows that altering the order of $(-600)$ and $900$ does lead to different results in the end.

*Discussions*

1. **Mathematical laws or not?** $a$, $b$ and $c$ are positive numbers. For each of the following, determine if the formula is a mathematical law or not. If it is a mathematical law, find out what it's called and explain why it works. If not, show an example of how it fails.

   a) $(a + b)^2 = a^2 + b^2$.

b) $(a+b)^2 = a^2 + 2ab + b^2$.

c) $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

d) $\sqrt{a \times b} = \sqrt{a} \times \sqrt{b}$.

e) $\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}$.

f) $a^b = b^a$.

g) $(a+b)+c = a+(b+c)$.

2. **Pancake batter.** A pancake recipe requires mixing $1$ cup of pancake mix and $0.75$ cup of water to make the batter. Does this recipe make $1.75$ cups of batter? Explain your answer. (Experiment in the kitchen if you want to.)

3. **The 1-plus-1-unequal-to-2 story.** Make your own story of $1+1 \neq 2$. Do you think it can be used to discredit the normally accepted math of $1+1=2$? Share your thoughts.

4. **Tweak another math you know.** Mathematical laws and formulas work under their assumptions. Tweak another math by considering a situation where the math doesn't work. Explain what you do and how you do it.

*Chapter 2*

# *Probability in Your Life*

## 2.1 What is Probability

The probability of something (like an event) is a number that measures how likely it would happen. This number is between 0 and 1.

- The larger the probability is for something, the more likely it would happen. For example, an event with 0.5 probability is more likely to happen than another event with 0.2 probability.
- Some people prefer to use percentages. In that way, the probability is between 0% and 100%.
- 0 (0%) means impossible to happen and 1 (100%) means absolutely happening.

### *Frequency As Probability*

People get a sense of *how likely* something happens by observing how *frequently* it happens. This is done by *looking at a large data*, or *conducting a large experiment or survey.* For example, during the 2019–2020 NBA Season, LA Lakers player LeBron James made 239 free throws out of 343 attempts. The frequency is $239/343 = 69.7\%$, which can be viewed as the probability for him to make a free throw at his next attempt.

As for using large experiments, assume that a factory wants to know how good a production line is, so it randomly tests a thousand items produced on the line: 998 of them are good and 2 of them are defective. Then the probability for the production line to produce a good item is $998/1000 = 99.8\%$.

Why does it have to be large enough? Because when the observation is small, what you see may not be representative of the general likelihood for something to happen. This is the underlying idea of *the Law of Large*

*Numbers*, which says that the data needs to be large enough for the observed frequency to reflect the probability in reality.

*Example: M&M's Chocolates*

In a big bag of M&M's chocolates, there are 24% blue, 20% orange, 16% green, 14% yellow, 13% red and 13% brown. If you pick one chocolate randomly from the bag, the probability of getting an orange one is 24% because if you pick a lot of them, about 24% of them would be orange.

*Question.* Can we say that the probability of each color is $1/6$ because there are six colors? Why or why not?

## *Random Choices*

When someone makes a random choice, it means the person gives each of the possible options the same chance (or probability). For example, if Ben has five ties and he picks his ties randomly, each tie has a 20% probability to be chosen. The total probability of 100% is *evenly* divided by the five ties: $100\% \div 5 = 20\%$. Here are some more examples.

1. 72 employees enter a random drawing for an iPad. 30 of them are full-time and 42 are part-time. Each person has $1/72$ chance to win. The full-time employees collectively have $30/72$ chance to win and the part-time ones have $42/72$ chance to win.

2. The Powerball is an American lottery. There are 292,201,338 different combinations and only ONE winning combination. If the lottery is random (or fair), the probability for a ticket to have the winning combination is 1/292,201,338, or approximately one out of three hundred million chances.

*Not All Choices Are Random*

In reality, most events or choices in life are not random, so the idea of equal likelihood doesn't apply. Thus, we cannot give each outcome equal probability. For example, a professional football team may have three quarterbacks, but the starting position is usually assigned to the same person. Another example: You are taking the Math for Liberal Arts course and there are two possible outcomes: pass or fail. This doesn't mean that you only have $100\% \div 2 = 50\%$ chance of passing it.

## *Too Good (or Bad) to Be True*

The probability is small for something that rarely happens. When it happens, people may express disbelief: "*No way! That's too lucky (or too bad)!*"

It's possible that sometimes someone does get very lucky or unlucky, but an unlikely event can also be the result of unfairness, like cheating. The 1980 Pennsylvania Lottery Scandal is a good example: It was too good to be true, and it turned out to be the result of cheating.

*Example: Workplace Discrimination*

A company hires 20 new employees *with the same job description*, and three of them are black. The manager gives the three worst shifts to the three black employees. When accused of discrimination, the manager claims that the assignment is made *randomly*, and the three black employees are just unlucky. Does this sound reasonable?

Mathematics can calculate and show that there are 1142 different ways to assign the three worst shifts to the twenty employees, so the probability for the three black employees to randomly get these shifts is $1/1142 = 0.00088 < 0.1\%$. This probability is extremely small, making it highly unlikely to happen. This result suggests that the assignment is not random but intentional. In this case, it is a strong indication that there is racial discrimination in the work assignment.

Note that this analysis only works against the *randomness assumption*. For example, a company may have over 100 employees but the bathrooms are always cleaned by the two custodians. That doesn't mean the company discriminates against the custodians. It only means that the cleaning job is NOT randomly assigned to all employees.

## *Discussions*

1. **120% Sure?** Sometimes people can go above 100% when comparing quantities, like *this year he is making 120% of last year's salary*. But can you be more than 100% sure about something? Explain your answer.

2. **100% Free Throw Shooter.** In the 2019/2020 NBA Season, Chicago Bull's Mark Strus made one free throw out of one attempt, which means his free throw percentage is $1/1 = 100\%$. Does this mean that he has $100\%$ probability of making his next free throw? What's your take on this?

3. **What's riskier?** According to a study,

    - The probability of dying in a plane crash is 1/205,552,
    - The probability of dying while cycling is 1/4,050, and
    - The probability of dying in a car accident is 1/102.

Based on these numbers,

a) Which one do you THINK is riskier?
b) Which one do you FEEL is riskier?
c) If they are different, is there an explanation?

4. **50-50 Right?** During an interview, NBA legend Shaquille "Shaq" O'Neal was asked to predict the outcome of an upcoming basketball game. Shaq famously responded, "Well, you either win or lose... I guess it's 50-50." What's your take on this?

5. **Randomness and Equal Opportunity.** The Equal Employment Opportunity Act of 1972 address employment discrimination against African Americans and other minorities. What is your take on equal opportunity? Do you think there is a connection between equal opportunity and the random choices discussed in this math lesson? Please explain.

## 2.2 *The Conditional Probability*

### *Probability with Context*

Probability can be considered in a particular context or within a specific group. When it is within a context, we call the probability a *conditional probability* and the context a *condition*. For example, the president's overall approval rating is a general probability, and *the president's approval rating in Minnesota* is a conditional probability, with the condition being the people of Minnesota.

Another example: The probability of anyone being a Facebook user is a general probability, and the chance for any teenager to be a Facebook user is a conditional probability with the condition being teenagers.

### *Example: Nursing Students*

Assume that at University X, the student population can be broken down to six groups. These percentages are of the entire student population, with total percentage $100\%$.

| | Nursing | other |
|---|---|---|
| male | $1\%$ | $35\%$ |
| female | $7\%$ | $47\%$ |
| nonbinary | $1\%$ | $9\%$ |

1. If we randomly choose *a female student*, what is the probability that she is in nursing? First, the female students are $7\% + 47\% = 54\%$ of the entire university and the female nursing students are $7\%$ of the entire university. So among all the female students, the conditional probability for a student to be in nursing is $(7\%)/(54\%) = 13\%$.

| | Nursing | other | subtotal |
|---|---|---|---|
| **female** | **7%** | **47%** | **54%** |

2. If we choose among all the nursing students, what's the probability that the student is nonbinary? The condition is nursing major, which is $(1\%) + (7\%) + (1\%) = 9\%$ of all, while $1\%$ of all is nonbinary in Nursing. So the conditional probability of getting an nonbinary student among Nursing students is $(1\%)/(9\%) = 11\%$.

| | **Nursing** |
|---|---|
| male | **1%** |
| **female** | **7%** |
| nonbinary | **1%** |
| subtotal | **9%** |

*Conditional Probability vs General Probability*

*Favorable, unfavorable or independent.* When comparing the conditional probability with the general probability of an event, if the conditional probability is higher than the general probability, the condition is said to be *favorable* for the event. If the conditional probability is lower than the general probability, the condition is *unfavorable* for the event. And if the conditional probability is the same as the general probability, the event is called *independent* of the condition.

*More likely than usual vs more likely than not.* When something is more likely than not, it has over $50\%$ probability. (Why?) On the other hand, a favorable condition results in a probability that is *higher than usual*, which may or may not be higher than $50\%$. FOr example, *clutch hitters* in baseball refer to those players who hit better than usual when there is a scoring opportunity. So, the scoring opportunity is a favorable condition for this player. It doesn't mean they hit at $50\%$ or higher.

*Unfavorable but not a bad thing.* Keeping one's cholesterol levels in the normal range can lower the risk of stroke. So, for the probability of getting a stroke, keeping normal cholesterol levels is an unfavorable condition because it *reduces* the probability. (It is a favorable thing for one's health, though.)

Back to the example of nursing students at University X:

| | Nursing | other |
|---|---|---|
| male | 1% | 35% |
| female | 7% | 47% |
| nonbinary | 1% | 9% |

The general probability for a student to be in nursing is $(1\%) + (7\%) + (1\%) = 9\%$, and the conditional probability for a female student to be in nursing is 13%, which is higher than the general probability. So, being female is a favorable condition to being in nursing. In common language, we say that *you are more likely to find a nursing student among the female students.*

*Favorability is mutual.* Mathematics can prove that, when a condition is favorable to an event, it is true vice versa. Using the nursing students in University X for example: you are more likely to find a nursing student amongst female students, and you are also more likely to find a female student amongst nursing students. The same is true for unfavorable or independent conditions.

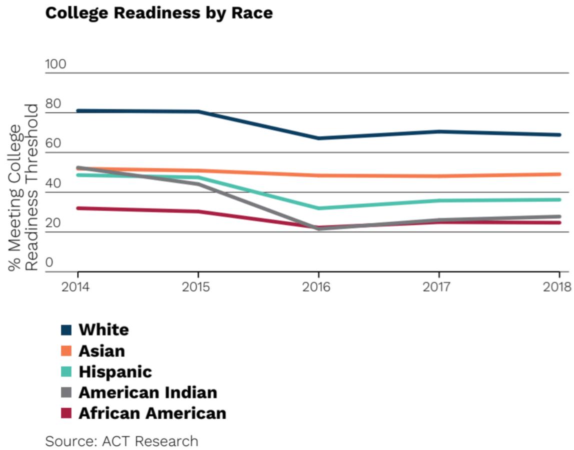
*Conditional Probability and Stereotypes*

According to https://nces.ed.gov, in 2015-2016, of all the US graduating college students, 18% are in the STEM fields (Science, Technology, Engineering and Math), while of all the US graduating Asian college students, 33% are in the STEM fields. So, an Asian student is more likely to have a STEM degree *than college students in general.* On the other hand, this conditional probability is still under 50%, so we shouldn't assume the stereotype of Asian students being in STEM fields because less than half of them are.

*Example: Achievement Gap in Minnesota*

Reference: A Statewide Crisis: Minnesota's Education Achievement Gaps by Federal Reserve Bank of Minneapolis

Minnesota consistently ranks among the highest-performing states when it comes to education quality. However, according to the article, there are persistent, significant gaps in the achievement of certain groups of Minnesota students as compared to others. Take college readiness, for example. The probability of college readiness is different under different conditions (of racial identity). This means *college readiness is NOT independent of race.*




*Discussions*

1. **Bias in real life.** The bias towards a person or a group can sometimes be captured by a favorable/unfavorable conditional probability. Give an example and explain what happens in it.

2. **Favorable, independent, or unfavorable conditions.** Find an example for each of the following. Be specific about the event and the condition.

   a) A conditional probability with a favorable condition.

   b) A conditional probability with an unfavorable condition.

   c) A conditional probability with an independent condition.

   ***Sample.*** *The event is "contracting lung cancer" and the condition is "having second-hand smoke". The condition is in favor of the event because those who have second-hand smoke have higher probability of getting lung cancer than the overall population, according to American Cancer Society.*

3. **A lot?** Someone said to you that *a lot of tall people are good at playing basketball.* Does this mean that a really large number of tall people are good at basketball, or that being tall is favorable for playing basketball well? Give your take on this, and then find a similar example.

4. **College Readiness by Race.** Read the whole article cited in the lesson about the racial gap in college readiness. State and comment on one evidence from this report. Please refererence the page number.

## 2.3 *COVID-19: Should Everyone Get Tested?*

**Disclaimer.** This is a mathematical model that makes more sense at the beginning of the COVID-19 pandemic with the assumption that there is a small percentage of infected individuals and a good tracking system to control the spread. This model may or may not apply to current day United States or any other country.

### *COVID-19 in New York*

The State of New York was hit the hardest by the pandemic at the beginning of USA's COVID-19 outbreak. On March 22, 2020, the state recorded 15,885 total number cases, with 15,276 active cases, 209 total death and 5,440 daily new cases. At the time, the state (and the entire nation) lacked the capacity to test all people with COVID-like symptoms. While the nation worked on getting enough testing, some people suggested that it would be great if the government would test *everyone*. Is that a good idea? Let's analyze this with mathematics.

### *Background Information*

In March 2020, the population of New York State was approximately 20 million, the COVID-19 infection rate was approximately $2\%$, and the COVID-19 testing had a $6\%$ false-positive rate and a $1\%$ false-negative rate.

Under these assumptions, $2\%$ of the people in the State of New York had COVID-19, or $(20\,\text{million}) \times (2\%) =$ 400,000 people. On the other hand, $98\%$ did not have COVID-19, or $(20\,\text{million}) \times (98\%) =$19,600,000 people.

### *Analysis*

Suppose all 20 million residents took the test.

1. Of the 19,600,000 people without COVID-19, 6% would get false-positives, which would be (19,600,000)×(6%)=1,176,000 people. 94% of them would get true-negative results, which would be (19,600,000)×(94%)=18,424,000 people.

2. Of the 400,000 people with COVID-19, 1% would get false-negatives: (400,000)×(1%)=4,000 people. 99% would get true-positive results:(400,000)×(99%)=396,000 people.

Sort this information into this table:

| | Testing (+) | Testing (-) |
|---|---|---|
| Not having COVID | 1,176,000 (false-positive) | 18,424,000 |
| Having COVID | 396,000 | 4,000 (false-negative) |

Looking at all the positive testing results: In total, there would be 1,176,000+396,000=1,572,000 of them, but the majority of them (1,176,000) were false-positives. If these folks went to emergency rooms or made a few panic calls, it would put unnecessary stress on the medical resources and social stability. And the 396,000 people with COVID-19 were at the minority, which makes it hard for government to identify and follow up with them.

*Remarks.*

1. The big problem of universal testing is the large number of false-positive cases that dilute the real COVID-19 cases. This problem can be reduced by prioritizing testing on those who have symptoms or were exposed to confirmed cases.

2. The false-negatives tell infected people that they are okay, but their number is small enough to re-test.

3. When the infection rate is higher, there will be fewer uninfected people, which results in fewer false-positive cases. But there will be more infected people and therefore more false-negative cases.

4. This analysis shows that, even when a medical testing is fairly accurate, if someone gets a positive testing result for an unfavorable disease, they should always test again to be sure.

*Discussion*

In the article about COVID testing, we identified the main problem as the large number of false-positives.

1. Is there a circumstance when we won't get a large number of false-positives?

2. Does this mean that we should never try to test *everyone*?

What are your thoughts?

## 2.4 Video: Should Computers Run the World?

Watch this documentary on YouTube: Should Computers Run the World? - with Hannah Fry

*Discussions*

1. **Sensitive vs specific.** What is your take on the video's comments on the two elements of algorithm: being sensitive and being specific?

2. **Summary of the video.** State and explain at least two takeaways you get from the video.

*Chapter 3*

# *Lengths and Areas*

## 3.1 Video: Hunting the Hidden Dimension

Watch this documentary: Hunting the Hidden Dimension

*Discussions*

1. **Ideas in fractal geometry.** Describe an idea or a concept in fractal geometry that impresses you most and explain why. Please refer to where it is at the video.
   ***Sample.***

   - *Concept: How iteration is used to generate fractals*
   - *Reference: 4:55 in the video*
   - *Comment: Blah blah blah.*

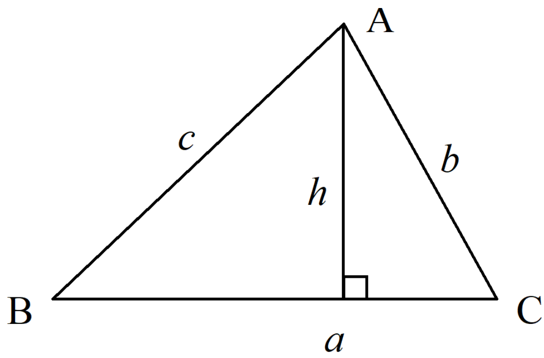
2. **Applications of fractals in technology.** Describe an application of fractals in technology that impresses you most and explain why. Please refer to where it is at the video.

3. **Fractal geometry in natural science.** Describe a fractal geometry in natural science that impresses you most and explain why. Please make reference to where it is at the video.

4. **Summary of the video.** Name three things you learned from this video. Present and describe them clearly.

## 3.2 The Areas of Triangles

Most people know from elementary or middle school math that, the area of a triangle is "base $\times$ height divided by 2". In other words, $A = \frac{bh}{2}$.




In high school trigonometry, there is a second formula. If we are able to measure angle $A$, then the area can also be calculated by $\frac{bc\sin A}{2}$ where $\sin A$ is the sine function of the angle $A$.

These two formulas, however, are not always suitable for practical uses. The first formula $A = \frac{bh}{2}$ requires us to draw and measure the height $h$, and the second formula requires us to measure the angle $A$ and find out about its sine value. Both requirements would be difficult if the triangle is very large (like a lake), or if the inside of the triangle cannot be accessed or drawn on (like a piece of cheese).

Wouldn't it be nice if all we have to do is measure the three sides?

## *Heron's Area Formula for Triangles*

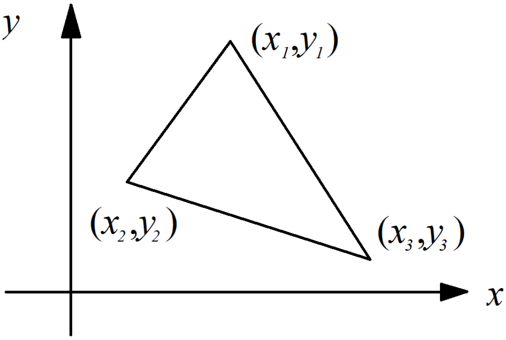
Heron of Alexandria (c. 10 AD – c. 70 AD) was a Greek mathematician and engineer who was active in his native city of Alexandria, Roman Egypt. Heron is most famous for inventing the formula that calculates the area of a triangle from the lengths of its three sides. If the three sides are $a$, $b$ and $c$, then the area is $\frac{\sqrt{(a+b+c)(a+b-c)(b+c-a)(c+a-b)}}{4}$. For example, if the three sides of a triangle are 10, 12 and 15, then no other measuring is needed: the area is

$$\frac{\sqrt{37 \cdot 7 \cdot 17 \cdot 13}}{4} = 59.81.$$

Some people call Heron's formula *Hero's formula*. Given how powerful this formula is, yes, Heron is truly a hero of mathematics.

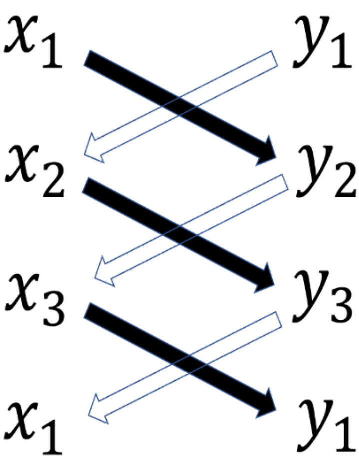
## *The Shoelace Formula for Triangles*

If a triangle is drawn on an $xy$-plane, then all three vertices have $xy$-coordinates. Assume that the three vertices are $(x_1, y_1)$, $(x_2, y_2)$ and $(x_3, y_3)$.




Then the area of the triangle is $\frac{|x_1y_2-y_1x_2+x_2y_3-y_2x_3+x_3y_1-y_3x_1|}{2}$.

(| | is the absolute value.) This formula does look complicated, but someone found a way to memorize the stuff inside | |:




where we cross multiply like the laces, add the products obtained from ↘ and subtract the products obtained from ↙.

This formula is extremely powerful because all you need are the coordinates of the vertices. For example, if the three vertices are $(1, 0)$, $(2, 5)$ and $(-1, 3)$, then the area is $\frac{|1\cdot5-2\cdot0+2\cdot3-(-1)\cdot5+(-1)\cdot0-1\cdot3|}{2} = 6.5$.

### *Discussions*

1. **The limitation of the traditional area formulas.** Give another example where the two formulas $\frac{bh}{2}$ and $\frac{bc\sin A}{2}$ are not suitable for practical uses.
2. **What is area?** In your own words, explain the idea of area to (a) a third grader and (b) an alien from a different planet.

3. **Other area formulas in geometry.** What other area formulas do you know? Parallelogram? Rhombus? Rectangles? Describe the formula and give an example of how to use it.

4. **Heron's formula.** Heron's formula only requires measuring the three sides. Think of a situation when Heron's formula is more useful than other area formulas.

5. **Practicing Heron's formula.** Calculate the area of the triangle from the three given sides.

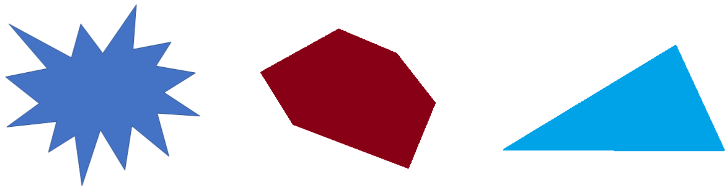
   a) 5, 7, and 9.
   
   b) 10, 13, and 13.
   
   c) 1, 2, and 3. Can this be calculated? What's going on?
   
   d) 5, 5, and 11. Can this be calculated? What's going on?

6. **Practicing the shoelace formula.** Calculate the areas of these triangles with given vertices.

   a) $(0, 1)$, $(3, -2)$ and $(1, 0)$.
   
   b) $(1, 2)$, $(-2, 1)$ and $(3, 3)$.
   
   c) $(0, 0)$, $(3, 5)$ and $(-2, 6)$.

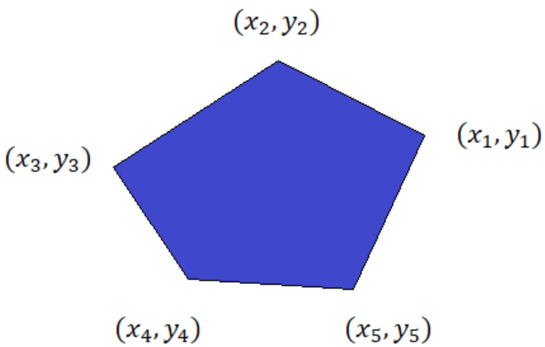
## *3.3 How Large is Bde Maka Ska?*

A polygon is a plane figure that is described by a finite number of straight line segments connected to form a closed area. Triangles, squares, trapezoids, pentagons, hexagons, and octagons are all polygons, so are the shapes with parts that are *caved in*, like stars.



## *The Shoelace Formula for Polygons*

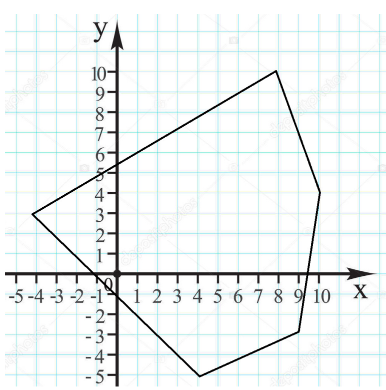
The shoelace formula applies for any polygons, as long as we set up the the vertices in order, like with the triangles. For example, for a pentagon




the area is $\frac{1}{2}\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \\ x_5 & y_5 \\ x_1 & y_1 \end{vmatrix}$, using the same shoelace pattern as the one for triangles.

### *Example: Area of a Pentagon*

We wish to find the area of this pentagon with vertices $(10, 4)$, $(8, 10)$, $(-4, 3)$, $(4, -5)$ and $(9, -3)$.

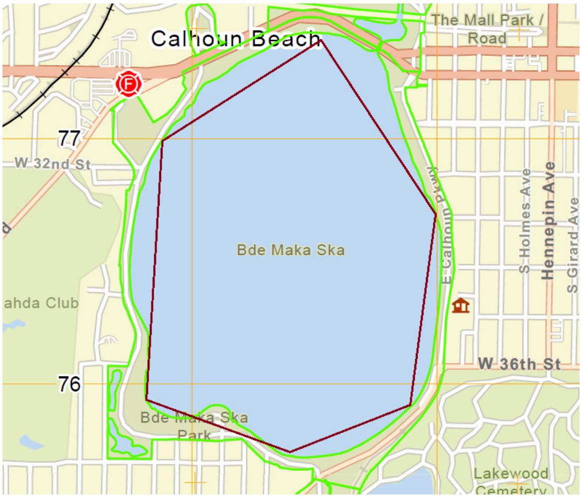


Simply set up the shoelace formula: $\frac{1}{2}\begin{vmatrix} 10 & 4 \\ 8 & 10 \\ -4 & 3 \\ 4 & -5 \\ 9 & -3 \\ 10 & 4 \end{vmatrix} = 119.5.$

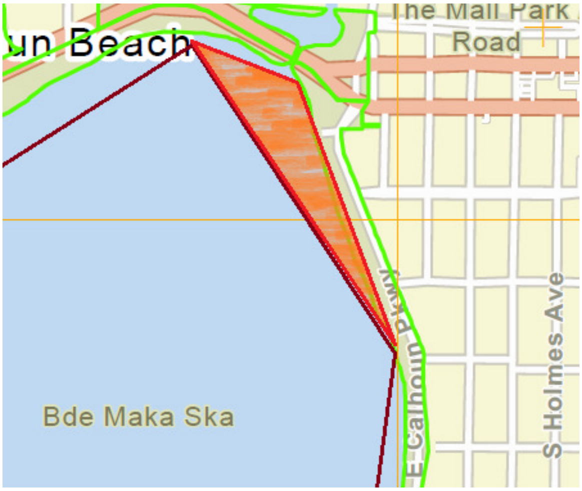
### *How Large Is Bde Maka Ska?*

Bde Maka Ska is the largest lake in Minneapolis, Minnesota, United States, and part of the city's Chain of Lakes. Surrounded by city park land and circled by bike and walking trails, it is popular for many outdoor activities. (See: Minneapolis Park and Recreation.)

How Do You Measure the Area of Bde Maka Ska? Let's begin with its map. Note that the grids are 1 kilometer apart. To estimate the area of the lake, we mark a few points in the lake shore and draw a polygon that roughly represents the lake. Notice that we only use 6 vertices here. This estimate is very rough, but the area of this polygon can be quickly calculated via the shoelace formula.




This is a rather simple estimate, but we can improve this result by going after the local regions that have more error. For example, the northeast corner. Once zoomed in, we fit in another polygon (in this case a triangle), which provides good improvement:




*Why is this a good approach?*

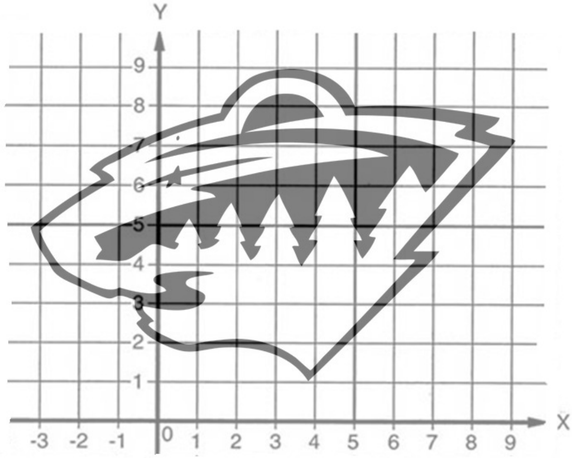
Quick to calculate. It is relatively easy to calculate the area using the shoelace formula (as long as we figure out the coordinates). Easy to improve. This method allows us to improve the result in efficiently and strategically. For example, I would choose to work on the northeast corner before the southwest corner, where there is less error.

*Getting the Job Done*

To put this approach to work and really get the final answer, we just need the coordinates of the vertices. Of course we can get a map that has fine grids for us to mark the points by hand. Better yet, we can get the GPS coordinates of those points (from an online map like Google Map), convert them into $xy$-coordinates, and then apply the shoelace formula. This conversion is well studied and commonly used in math and geography, but does require additional work in math. We shall skip it for now.

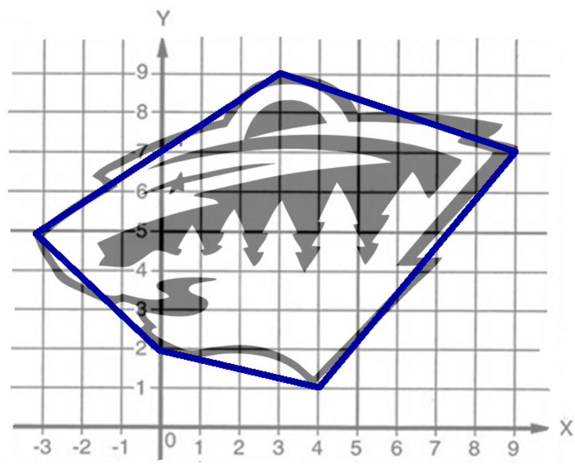
*Discussions*

**Minnesota Wild Logo.** The Minnesota Wild is a professional hockey team. A local sports bar orders a team logo for display. Below is the design, with the grids being 1 foot apart. We wish to find out the area of this display.




1. How would you calculate this area? You may use any method you know, but specify your method and explain how you used it to get your answer.

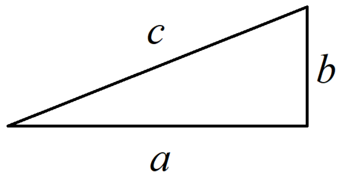
2. Find an estimate of the area by considering this pentagon and using the shoelace formula.




## 3.4 *How Long Is the Shoreline of Bde Maka Ska?*

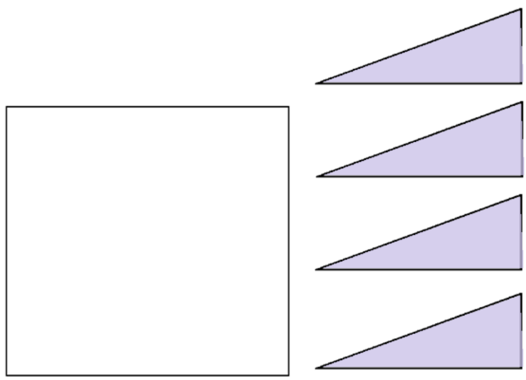
### *The Pythagorean Theorem*

For a right triangle with perpendicular sides $a$ and $b$ and the hypotenuse $c$, the three sides have a special relation: $a^2 + b^2 = c^2$. This is the famous *Pythagorean Theorem.*

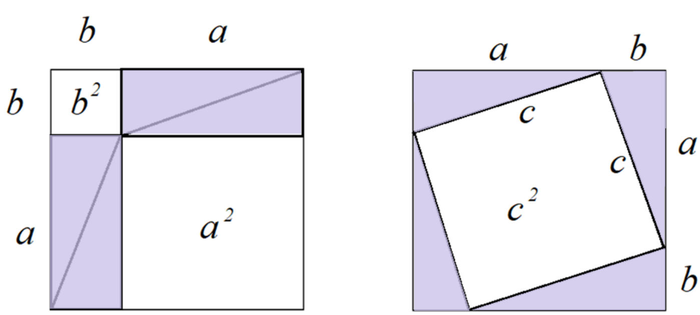



*Proof of the Pythagorean Theorem*

Mathematical theorems require proof. Here is one for the Pythagorean Theorem. Consider four right-triangle-shaped rugs and a square floor of side length $a + b$ .



If we place the rugs one way, the uncovered bare floor area is $a^2 + b^2$. But if we rearrange them in a second way, the uncovered bare floor area is $c^2$.




Since the uncovered floor area should be the same no matter how we place the rugs, $a^2 + b^2 = c^2$. □

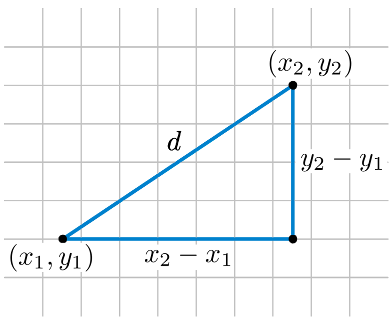
*The Pythagorean Triples*

Pythagoras and his followers believed that the universe is created based on mathematical rules and *whole numbers*. Ironically, they had difficulty finding whole numbers for hypotenuses, although they did discover a few. If the three sides of a right triangle are all whole numbers, we call those three numbers a *Pythagorean triple*. For example, $(3, 4, 5)$, $(5, 20, 21)$ and $(8, 15, 17)$ are all Pythagorean triples.

The 3-4-5 triangle is also called the *carpenter's right triangle*. To get a right angle, a carpenter can make a right triangle easily by using three pieces of wood that are 3-, 4- and 5-inches long. This is easy to do and yet very accurate. The Egyptians and the Chinese also discovered this particular use of 3-4-5 triangles at the early stages of their civilizations. Next time you wonder how the Egyptians constructed the pyramids with such accuracy, remember that the 3-4-5 triangles has a role in it.

*The Length of Line Segments*

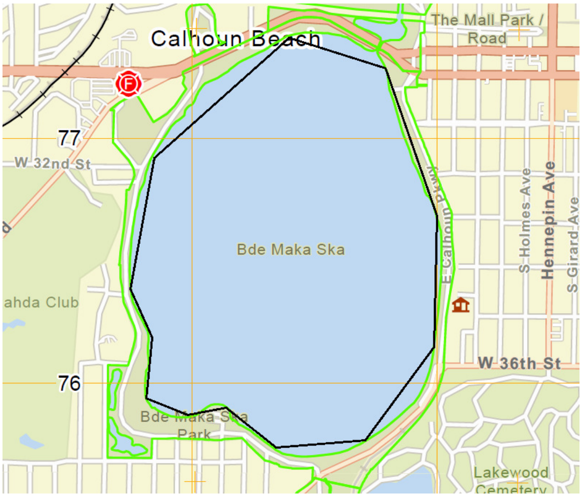
There is a formula to calculate the distance between two points on the $xy$-plane. This distance is also the length of the line segment connecting the two points.




Let's call this line segment $d$, which is the hypotenuse of the right triangle. Also observe that the other two sides are $(x_2 - x_1)$ and $y_2 - y_2$. By the Pythagorean Theorem, $(x_2 - x_1)^2 + (y_2 - y_1)^2 = d^2$. So $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.

*Measuring the Shoreline of Bda Maka Ska*

To estimate the length of the shoreline, we mark a few points and connect them to a polygon.




Then we calculate each line segment using the length formula mentioned above.

*Discussions*

1. **Pythagoras.** Greek mathematician Pythagoras was among the greatest mathematicians of all time. Find out more about Pythagoras, his school of thoughts, his discoveries, and his influence on the world civilizations. Name two or more things that impress you.

2. **The closer you look the longer the shoreline?** The video of the week, *Hunting the Hidden Dimension*, discusses about the problem of measuring the shoreline of Great Britain. Does that apply to measuring the shoreline of Bde Maka Ska? Explain what you think and how you would address this issue.

*Chapter 4*

# Logic Part 1 Making Sense with Logic

## 4.1 Introduction

When people reason with each other, they may say: *because of (A) and (B), therefore (C).* This sentence pattern is a structure called an **argument**, where (A) and (B) are called *premises* and (C) is called the *conclusion* of the argument. The process of going from the premises to the conclusion is called an *inference*.

Logic is the analysis and appraisal of arguments. When someone makes a valid argument, we say they are *being logical* or *making sense*.

### *Propositions*

The most basic element of logic is a proposition. A proposition is a claim that can be verified as *true or false but not both or neither*. For example,

- Honeycrisp apples originated in Minnesota.
- The average price of gas today is $2.99/gallon.

We use letters to denote propositions, $0$ to denote FALSE and $1$ to denote TRUE. For example, $p$ : "*Honeycrisp apples originated in Minnesota*". This proposition is indeed true. (See: Wikipedia.) So $p = 1$.

Whether a proposition is true or false may depend on the circumstances or the person who claims it. For example, for someone born in 2000, "*I am* $21$ *years old and I can legally buy alcohol*" is false in 2020, but true in the year 2021.

*The Non-Propositions*

Not all sentences are propositions. For example, sentences that don't claim anything are not propositions. (Like "*How are you doing?*" or "*Have you watched that movie?*".) Another type of non-proposition are the paradoxes, or self-contradictions. We will talk about paradoxes later in this chapter.

*The Word AND*

Two propositions can be connected with the word AND. The combined proposition would be true *only when both propositions are true.* We use the notation $\wedge$ for the word AND.

For example, to apply for a Minneapolis College Foundation Scholarship, someone needs to (*be a student at Minneapolis College*) AND (*have at least 2.3 GPA*). If I am interested in applying for the scholarship, there are 4 possible scenarios:

1. $1 \wedge 1 = 1$: I am a student and have a qualifying GPA. I **can** apply.
2. $1 \wedge 0 = 0$: I am a student at the College. I don't have a qualifying GPA. I **cannot** apply.
3. $0 \wedge 1 = 0$: I am not a student at the College. I have a qualifying GPA. I **cannot** apply.
4. $0 \wedge 0 = 0$: I don't meet either criterion. I **cannot** apply.

Mathematicians often record these results with a table (which is called a truth table):

| $p$ | $q$ | $p \wedge q$ |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |

*The Word OR*

Two propositions can be connected with the word OR, and the combined proposition would be true as long as one proposition is true. We use the notation $\vee$ for the word OR.

For example, to enter the security check at the airport, a passenger needs to bring (*a valid passport*) OR (*a REAL ID*). For a passenger trying to enter the security check, there are 4 possible scenarios:

- $1 \vee 1 = 1$: Having both the passport and the REAL ID, the passenger may enter.

- $1 \vee 0 = 1$: Having the passport but not the REAL ID, the passenger may enter.
- $0 \vee 1 = 1$: Having the REAL ID but not the passport, the passenger may enter.
- $0 \vee 0 = 0$: Having neither the passport nor the REAL ID, the passenger may NOT enter.

These results can be recorded with this truth table:

| $p$ | $q$ | $p \vee q$ |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 1 |
| 0 | 1 | 1 |
| 0 | 0 | 0 |

*The Word NOT*

We can phrase the opposite of a proposition by adding the word "NOT" to it (with all the necessary grammatical adjustments). The notation is $\neg$. So $\neg 0 = 1$ and $\neg 1 = 0$.

## *The Dichotomy in Logic*

A dichotomy separates all possible cases into two categories. For example, *everything purple* vs *everything not purple.* The dichotomy allows us to use a simple but effective reasoning pattern: If it is not one, it must be the other. In logic, propositions are either true or false. This is a dichotomy.

*Example: Who Wants Cream?*

Three mathematicians, Javier, Ben, and Maria, go into a coffee shop and each orders a cup of coffee.

- The barista: Does anyone want cream?
- Javier: I don't know.
- Ben: I don't know.
- Maria: Yes.

So, who wants cream and who doesn't want cream?

*Solution.* The question here is "*Does anyone want cream*".

1. Javier either wants cream or not. If Javier wants cream, he would answer YES (someone wants cream). However, he does not answer yes. So, Javier does NOT want cream.

2. Ben logically deduces that Javier doesn't want cream. Again if he wants cream he would answer YES. However, he does not answer yes. So, Ben does NOT want cream.

3. Maria logically deduces that neither Javier nor Ben wants cream. She then answers YES. So Maria wants cream.

The final answer: Javier and Ben don't want cream but Maria wants cream.

### *Discussions*

1. **Examples of Dichotomy.** A dichotomy separates all the things in discussion into two categories without overlapping or leaving anything out. Are these dichotomies? Why or why not? Make your own example of dichotomy.

   a) She loves me vs she hates me.

   b) She loves me vs she doesn't love me.

   c) Republicans vs Democrats.

2. **Dichotomy Way of Thinking.** Some people tend to use dichotomy as a thinking pattern of: "*it's either this or that*".

   a) Given an example.

   b) What do you call them or how do you describe this behavior? (For example, this is a black-and-white kind of person.)

   c) Do you agree with this type of thinking? Why or why not?

3. **Who Wants Sugar?** On another day, Javier, Ben and Maria go to the coffee shop again, and this time there is a different barista.

   - The barista: Does EVERYONE want sugar?
   - Javier: I don't know.
   - Ben: I don't know.
   - Maria: No.

   So, who wants sugar and who doesn't? Explain your answer.

## 4.2 Tautologies, Contradictions, and Paradoxes

### Tautologies

A tautology is a proposition that's always true. The most well-known tautology is $p \vee (\neg p)$: $p$ and $\neg p$ are always one true and one false, and when they are connected with OR, the result is always true. For example, "*I live in Minneapolis OR I don't live in Minneapolis*" is a tautology.

There are other tautologies that have more complicated structures. Since tautologies are true *by structure*, it doesn't require a reality check. In other words, tautologies have zero "specificity".

### Contradictions

A contradiction is a proposition that's always false. A well-known contradiction is $p \wedge (\neg p)$: $p$ and $\neg p$ are always one true and one false, and when connected with AND, the result is always false. For example, "*I am a Minnesota Vikings fan AND I am not a Minnesota Vikings fan*" is a contradiction. Contradictions, like tautologies, have zero "specificity": it is guaranteed by structure to be false and does not require fact-checking.

### Paradoxes

A proposition is required to be either true or false. Can there be a sentence that is neither? Such a sentence is called a paradox or self-contradiction. Consider the sentence "*This is a false sentence.*"

- If "*this is a false sentence*" is true, then what the sentence says is true, which makes it false.
- If "*this is a false sentence*" is false, then the opposite of what it says is true, which makes it true.

In both cases, "*This is a false sentence*" can't be determined as true or false. So it is NOT a proposition.

### Shield and Spear

The shield and spear paradox is a paradox in non-western thought, found in a story in the 3rd century BC Chinese philosophical book Han Feizi. In the story, a man was trying to sell a spear and a shield. When asked how good his spear was, he said that his spear could pierce any shield. Then, when asked how good his shield was, he said that it could defend from all spear attacks. Then one person asked him what would happen if he were to take his spear to strike his shield; the seller could not answer.

This led to the idiom of "zixiang maodun", meaning "self-contradictory". (Read the entire article on Wikipedia.)

*Knights and Knaves*

Knights and Knaves is a type of logic puzzle where some characters can only answer questions truthfully, and others only falsely. The name was coined by Raymond Smullyan. (See the related article on Wikipedia.)

On the Island of Knights and Knaves, everyone is either a knight or a knave. Knights always tell the truth and knaves always tell lies. If you ask someone "Are you a knight?", you will always get the answer yes. The reason?

- If this person is a knight, the answer is yes and (s)he will answer truthfully and say YES.
- If this person is a knave, the answer is no but a knave will answer falsely and say YES.

*Discussions*

1. **A Logical Dilemma.** There is a dilemma in answering the question: *Can the Almighty God create something that He cannot break?* Give your account based on logic, not your personal religious belief.

2. **More Paradoxes.** We talk about the Shield and Spear Paradox in the lesson. Do some research and find another paradox. In your own words,
   a) Describe/state the paradox,
   b) Explain why this is a paradox, and
   c) Remember to refer to your source of info.

3. **Knights and Knaves.** On the Island of Knights and Knaves, everyone is either a knight or a knave. Knights always tell the truth and knaves always tell lies.
   a) Find a question that knights and knaves will both answer YES.
   b) Find a question that knights and knaves will both answer NO.
   c) Find a question that knights will answer YES and knaves will answer NO.
   d) Find a question that knights will answer NO and knaves will answer YES.

4. **More on Knights and Knaves.** On the Island of Knights and Knaves, everyone is either a knight or a knave. Knights always tell the truth and knaves always tell lies. You meet with two islanders together. Not knowing if they are knights, knaves, or one of each, you ask them: "*Are you two of the same kind*?"

   Discuss how this question helps you find out their identities.

   a) What happens if both are knights?

   b) What happens if both are knaves?

   c) What happens if one of them is a knight and the other is a knave?

## 4.3 Conditional Propositions

*... I don't have much money, but boy, if I did*
*I'd buy a big house where we both could live...*
*– Your Song by Elton John (lyrics by Bernie Taupin)*

In the song, Elton John sings that if he had much money, then he'd buy a big house where he and his lover can both live. This type of proposition is called a conditional proposition.

**The Conditional Proposition** has the structure of "if $p$ then $q$". It is denoted by $p \to q$, where $p$ is called the *condition* and $q$ is called the *consequent*. A conditional proposition is like a *promise*: It is committed to the consequent $q$ as long as the condition $p$ is satisfied. If the condition is satisfied but the consequent is not fulfilled, the proposition is deemed to be *false*. Otherwise the proposition is considered *true* because the promise is not broken.

In Elton John's case, there are four possible scenarios:

1. $1 \to 1 = 1$: Elton John DOES have much money and he DOES buy a big house where he and his lover can both live. **He keeps his promise**, so the proposition is **true**.

2. $1 \to 0 = 0$: Elton John DOES have much money but he DOESN'T buy a big house where he and his lover can both live. **He does not keep his promise**, so the proposition is **false**.

3. $0 \to 1 = 1$: Elton John DOESN'T have much money but he DOES buy a big house where he and his lover can both live. (Perhaps

he can get a mortgage?) **He does not break his promise**, so the proposition is deemed **true**.

4. $0 \to 0 = 1$: Similarly, he does not break his promise, so the proposition is deemed **true**.

Overall, we can record these in the truth table:

| $p$ | $q$ | $p \to q$ |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 0 | 0 |
| 0 | 1 | 1 |
| 0 | 0 | 1 |

### *Condition vs Causation*

The conditional proposition $p \to q$ promises the consequent $q$ under the condition $p$. Does that mean $p$ directly causes $q$? Not always.

One possible situation is, $p$ may be just an indicator of $q$. For example, if *there are sirens from fire engines* then *there is a fire somewhere.* Clearly the sirens do not "cause" the fire.

With $p \to q$ we also cannot say $q$ happens because of $p$. Even if $p$ is a cause of $q$, *it may be one of the possible causes of (q).* For example, some students take bus to school and some students drive luxury cars to school. It is true that if the students have luxury cars then they can get to school. However, it would be very misleading say that the students can go to school is *because* they have luxury cars.

### *Discussions*

1. **Computing with Logic Symbols.** Calculate the following.

   a) $\neg(1 \wedge 0) \vee 0$

   b) $(1 \to 0) \to 0$

   c) $(\neg 1) \wedge (\neg 0)$

2. **Causation.** In a June 15, 2020 tweet, President Trump said testing “makes us look bad.” And in an interview with Fox News that aired on Sunday June 14, 2020, Trump could not have been clearer: “Cases are up because we have the best testing in the world and we have the most testing.”

   Is it logical to say having the best testing and the most testing is the reason (or the cause) for the number of COVID-19 cases to go up?

## 4.4 Videos: The Golden Ratio

Watch the following documentaries.

1. Golden Ratio = Mind Blown
2. The Golden Ratio and Fibonacci Sequence in Music

### Discussions

1. Name two things you learned from the videos. Please be specific.
2. Name one thing that is interesting enough for a casual conversation. ("Hey folks, I just learned this from my math class…"). Present it in an interesting enough way for such a conversation.

*Chapter 5*

# Logic Part 2 How to Argue It

## 5.1 Logical Equivalences

In algebra, two different algebraic expressions can give exactly the same results at all times. For example, no matter what we plug in for $x$ and $y$, $y + x + y$ and $x + 2y$ always have the same result. We say they are *equal*, and we write $y + x + y = x + 2y$.

Is there something similar in logic? Yes. Consider a proposition $p$ vs its *double negation* $\neg(\neg p)$:

- When $p = 1$, $\neg(\neg p) = \neg(\neg 1) = \neg(0) = 1$.
- When $p = 0$, $\neg(\neg p) = \neg(\neg 0) = \neg(1) = 0$.

Whatever value we have for $p$, $p$ and $\neg(\neg p)$ have the same value. We call $p$ and $\neg(\neg p)$ *equivalent*, and we use the notation $p \equiv \neg(\neg p)$.

When two proposition are (logically) equivalent, they can be viewed as the same. For example, saying *I like apples* is logically the same as saying *I don't dislike apples*, barring language-related preferences. When we say "*in other words*" in daily conversation, it means we are rephrasing a proposition by making an equivalent proposition.

The mathematical technique to analyze logical equivalences is beyond the scope of this lesson, but some other commonly used equivalences include: $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$ and $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$.

*Example of* $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$

You can pay for pizza delivery when you have (*cash* OR *credit card*). So you cannot pay for pizza delivery when you have (*no cash* AND *no credit card*). Note that the word OR is flipped to AND.

*Universal Propositions*

Propositions that talk about all things in general are called *universal propositions.* To show a universal proposition is false, it only takes one special case that doesn't work. This special case is called a *counterexample* to the universal proposition. For example, "*All roses are red*" is a universal proposition; to prove it false we can use a yellow rose as a counterexample.

In the English language, a proposition may indicate a universal sense without using the words *all* or every. It is the user's responsibility to carefully phrase a proposition whether (s)he means universally or not.

*Example: Minnesota's Achievement Gap*

Assume that you are a reporter covering Minnesota's educational policies. If you write "*Minnesota's K-12 schools have large achievement gaps*", this logically means that **all of** Minnesota's K-12 schools have large achievement gaps. Your opponents can attack this proposition by naming one counterexample. They just need to find one K-12 school that doesn't have a large achievement gap, and your proposition no longer stands. This kind of debate is unfruitful, but very logical.

In order to avoid such disputes, you should tailor the proposition and articulate it by saying "The *majority* of Minnesota's K-12 schools have large achievement gaps", or "*On average*, Minnesota's K-12 schools have larger achievement gaps than the rest of the nation."

## *Discussions*

1. **Tailoring your proposition.** In the example of describing Minnesota's achievement gap, we see how easy universal propositions can be overthrown. Find another example of a universal proposition, how it can be opposed, and how to fix it so that it would logically describe what you really mean.

2. **Equivalent Propositions.** The following pairs of propositions are equivalent. Make an example for each of them. Explain why they are equivalent.

   a) $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$.

   b) $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$.

   c) $p \equiv \neg(\neg p)$.

3. **Equivalent Forms of Conditional Propositions.** It is known that the three following propositions are equivalent. Make an example for $p$ and $q$ and phrase all three of them:

a) $p \to q$

b) $(\neg p) \lor q$

c) $(\neg q) \to (\neg p)$

Do they sound equivalent to you? Explain your answer.

***Sample.*** *$p$: "You don't stop" and $q$: "I will tell your supervisor". The three propositions become:*

*a) If you don't stop, then I will tell your supervisor.*

*b) Stop, or I will tell your supervisor.*

*c) If I don't tell your supervisor then (we can conclude) that you have stopped.*

.

## 5.2 Arguments

An *argument* is a structured collection of propositions. It contains a group of propositions called *premises* and a final proposition called the *conclusion*. An argument claims that *the conclusion is true whenever the premises are true*. If that is the case, we call the argument *valid*.

The validity of an argument is all about its structure. A valid argument *structurally guarantees the truth of the conclusion*, as long as the premises are true. When someone makes a valid argument, the only logical way to challenge it is to question the premises.

For example, consider this argument: *They are tall. They are rich. Therefore, they are professional basketball players.* All three propositions can be true for some people, but this is still not a valid argument because the truth of the premises (being tall and rich) gives no guarantee to the conclusion. We can find another person who is tall and rich but not an NBA player.

### Commonly Used Arguments

**The direct reasoning** is also known as *modus ponens* in Latin. It begins with a conditional proposition, confirms the condition, and concludes the consequent. It has this form:

$$\begin{array}{l} p \to q \\ \underline{p \qquad\quad} \\ \therefore q \end{array}$$

(The symbol $\therefore$ means *therefore.*)

For example,

If I am 18 then I can vote.
I am 18.
Therefore, I can vote.

**The indirect reasoning** is also known as *modus tollens* in Latin. It begins with a conditional proposition, has a negative consequent, and concludes a negative condition. It has this form:

$$\begin{array}{l} p \to q \\ \neg q \\ \hline \therefore \neg p \end{array}$$

For example,

If it rains, then I use an umbrella.
I don't use an umbrella.
Therefore, it does not rain.

**The transitive reasoning** argues in this form:

$$\begin{array}{l} p \to q \\ q \to r \\ \hline \therefore p \to r \end{array}$$

The transitive reasoning has two premises $p \to q$ and $q \to r$, and it concludes that $p$ can directly lead to $r$ like a chain reaction. For example, during the pandemic, "*if you live in Minneapolis, then you have free Wi-Fi*" and "*if you have Free Wi-Fi, then you can take online classes.*" Therefore, if you live in Minneapolis, then you can take online classes.

The chain of implications can be longer for the transitive reasoning. For example, with three premises $p \to q$, $q \to r$ and $r \to s$, we can link them all the way and conclude $p \to s$.

**The elimination reasoning** starts with two possibilities ($p$ or $q$), eliminates the first possibility ($\neg p$), and concludes the second and the only remaining possibility $q$. The format is:

$$\begin{array}{l} p \vee q \\ \neg p \\ \hline \therefore q \end{array}$$

For example, "*Mom or Dad will pick me up after school today*" and "*Mom can't come today.*" Therefore, Dad will pick me up today.

For the elimination reasoning, we can also begin with more than two possibilities and eliminate one or some of them. For example the format of eliminating two out of three possibilities:

$$\begin{array}{l} p \vee q \vee r \\ \neg p \\ \neg q \\ \hline \therefore r \end{array}$$

*Discussions*

1. **Arguments.** For each of the following arguments, what is the conclusion? Do you think the argument makes sense?

    a) The Minnesota Vikings played a game last Sunday. Since they did not win and they did not tie, they must have lost the game.

    b) My friend has a college degree and a high-paying job, so she must be happy.

    c) Club members can use the swimming pool. I am a club member, so I can use the swimming pool.

2. **Valid arguments.** Make an example for each of the following. One argument per thread per person.

    a) The direct reasoning.

    b) The indirect reasoning.

    c) The transitive reasoning.

    d) The elimination.

3. **Generalized elimination.** The original elimination argument has the premises $p \vee q$ and $\neg p$, which means there are 2 possibilities ($p$ or $q$) but $p$ is ruled out. Propose the format for the following and make an example.

    a) Eliminating 1 out of 3 possible options.

    b) Eliminating 2 out of 4 possible options.

    c) Eliminating 3 out of 4 possible options.

## 5.3 More about Arguments

### Logical Fallacies

Fallacies are invalid arguments. There are two invalid arguments that people often use by mistake: The *fallacy of the inverse* and the *fallacy of the converse.*

**The fallacy of the inverse** is also known as the inverse error. It has the format:

$$\begin{array}{l} p \to q \\ \neg p \\ \hline \therefore \neg q \end{array}$$

The argument says that in a conditional proposition $p \to q$, when we don't have the condition ($\neg p$), we would not have the consequent ($\neg q$). The argument is invalid because $p$ may be one of several possible causes of $q$. For example, if someone has a car then (s)he can get to school on time, but for those who don't have cars, we cannot conclude that they cannot get to school on time.

**The fallacy of the converse** is also known as the converse error. It has the format:

$$\begin{array}{l} p \to q \\ q \\ \hline \therefore p \end{array}$$

The argument says that in a conditional proposition $p \to q$, when have the consequent ($q$), we would confirm that we have the consequent ($p$). This is again invalid when $p$ is one of several possible causes of $q$. For example, if someone has a car then (s)he can get to school on time, but for those who get to school on time, we cannot conclude that they all have cars.

### Proof by Contradiction

The indirect reasoning (modus tollens) is a way to argue indirectly. Another broadly used method is *proof by contradiction.* The idea of proof by contradiction is that *a proposition is false if it leads to a contradiction.* So, to prove a proposition true, we instead show that its opposite is false. And this will be done by assuming the opposite to be true and logically deduce to something that contradicts an existing assumption or known truth. In other words, we show that *the opposite will lead to something wrong.*

*Example: New York Jets 2017*

The New York Jets (NYJ) are a professional football team in the National Football League (NFL). The New York Jets lost a football game on December 17, 2017 and their record dropped to 5 wins and 9 losses. There are 16 games in that season, so they only have two games left. You conclude that they would be a losing team for the season and their winning percentage would not get to 50% no matter what.

One explanation is that, *they need to win at least four games to get to 50% but they don't have four games*. If you choose to go this way, you have conducted a proof by contradiction, probably unknowingly.

First you assume the NYJ could get to $50\%$. (Assuming the opposite of what's to prove.) Then you deduce that the NYJ would need to win at least 4 more games. But there were only 2 games left in the season, which is a contradiction. So the assumption that NYJ could get to 50% must be false, meaning the team could NOT get to $50\%$.

*A Note of History*

The idea of proving something by showing the opposite wrong is commonly used by many civilizations. However, it was first conceptualized by the Greek mathematician Euclid of Alexandria (mid 400 B.C) who used logic extensively in mathematics and especially in geometry.

In his 13-volume book, the Elements, Euclid derived and proved all the mathematical discoveries until his time as the logical conclusions of a set of definitions and axioms. (Axioms are basic things we assume without proof.) *The Elements of Euclid* is among the most influential books in human history: its content is known to and studied by more people in the world than the content of any other book, including the Bible.

*Are Negative Political Campaigns Logical?*

Politicians sometimes spend time attacking their opponents by showing voters how wrong things would go if their opponents are elected. Is this logical?

Consider $p$: *voting for Candidate X* and $c$: *something going wrong*. The proof by contradiction method tells us that "*You should vote for Candidate X*" is logically equivalent to: "*If you don't vote for Candidate X then something will go wrong*". So, negative campaigning is absolutely logical. In a bipartisan competition, there is often only one other candidate, and the proposition becomes: If you vote for the other candidate then something will go wrong.

### *Discussions*

1. **Euclid and the Elements.** Research on Euclid and the Elements. What did you learn about them? What influence do they have on the world's many civilizations?

2. **The Inverse and Converse Errors.** Find a conditional proposition in a news article. Phrase its inverse error and converse error. Within the context, show why they are indeed errors.
   ***Sample.*** *According to MPR news on 8/28/2020, Minnesota recommends students returning to college... to get tested if they're having symptoms related to COVID-19. The Original proposition is: "If you are having symptoms, then you get tested."*

   - ***The inverse error****: If you don't have symptoms, then you don't get tested.*
   - ***The converse error****: If you get tested, then you have the symptoms.*

   *Both errors fail to apply to a person getting tested for other reasons such as the exposure to a confirmed case.*

## 5.4 *Video: Can We Trust Maths?*

Watch this documentary: Can We Trust Maths? - with Kit Yates

### *Discussions*

1. **28:29 of Video.** It was presented that:

   - Black people killed by police: 11%.
   - White people killed by police: 16%.
   - Black people killed by white people: 9%.
   - White people killed by white people: 81%.
   - White people killed by black people: 16%.
   - Black people killed by black people: 89%.

   What is your interpretation of these statistics? State and explain it.

2. **Video overview.**

   a) Name two things you learned from this video. Please be specific.

b) Name one thing that is interesting enough for a casual conversation. ("Hey folks, I just learned this from my math class…"). Do it in an interesting enough way for such a conversation and report back what your friends say.

*Chapter 6*

# *To Infinity... And Beyond!*

## *6.1 Large Numbers, Tiny Numbers*

According to US Census Bureau, there were 5303925 people living in Minnesota in 2010. We can talk about this number in the following ways:

1. **Using the numbers directly:** Minnesota's population is 5303925 in 2010.

2. **Adding the commas:** Minnesota's population is $5,303,925$. This is easier to read than 1) because we can quickly see that 5 is at the million's place. This approach also makes it easy to talk about 3):

3. **Counting by the million:** Minnesota's population is about 5.3 million.

*The Size of Numbers*

When talking about Minnesota's population, we can say that it is *in millions* or *best measured by the million* because its highest digit 5 is at the million's place:

$$\underbrace{5}_{\text{million's place}}, 3\ 0\ 3\ ,\ 9\ 2\ 5$$

In other words, the "size" of the number $5,303,925$ is the million. Since $5303925 = 5.303925 \times$ (one million), we can say Minnesota's population is about 5.3 million. In science, a million is conveniently recorded as $10^6$, so the number 5303925 is presented as $5.303925 \times 10^6$, which is the so-called *scientific notation* of the number. (For a short lesson on how to find the scientific notation of a number, see, for example, this Khan Academy video.)

The population of Minneapolis is 425403 in 2018. The highest digit 4 is at the hundred-thousand's place, so its size is a hundred thousand. Thus, we say 4.3 hundred thousand people live in Minneapolis, or hundreds of thousand people live in Minneapolis.

### *Talking about Tiny Numbers*

The diameter of the COVID-19 virus is between 0.00000006 and 0.00000014 meters. (See: https://www.news-medical.net) Let's talk about the tiny number 0.00000006:

$$0\,.\,0\,0\,0\,0\,0\,0\,0\qquad \underbrace{6}_{\text{(1/hundred million)'s place}}$$

so 0.00000006 is really 6 (hundred million)-th, or $6/100,000,000$. Therefore, the best size to measure up 0.00000006 is a hundred-million-th, or $1/100,000,000$.

Handling tiny numbers is similar to handling large numbers, except that we are working behind the decimal dot, and using negative exponents for 10. For example, $1/100,000,000 = 10^{-8}$, so the scientific notation of 0.00000006 is $6 \times 10^{-8}$ where $10^{-8}$ indicates the size of the number, 1 hundred-million-th.

## *Making Sense of Large Numbers*

Which of the following gives you a clear idea of the speed of light?

1. Light travels 670,600,000 miles per hour.
2. Light travels at 186,282 miles per second.
3. Light can travel around the Earth approximately 7.5 times in a second.

For many people, it is hard to grasp the concept of large numbers: After a certain size, they are all just very big, right? For example, a million might feel like a lot of money to most people, but they may not be proportionally impressed by a billion dollars (which is a thousand times a million dollars) or a trillion dollars (which is a million times a million dollars). We demonstrate a few ways to put large numbers in perspective and to better relate to them.

### *Converting to Rates*

Consider this example: Bill Gate's net worth increased by 16 billion in 2019. How much is 16 billion per year?

1. There are 365 days in 2019. On average, Bill Gates made $\$16,000,000,000/365 = \$43,835,61$, or roughly 44 million dollars per day.

2. There are 525,600 minutes in 2019. On a per minute average, he made $\$16,000,000,000/525,600 = \$30,441$, or approximately 30 thousand dollars per minute.

*Comparing with Another Large Number*

The speed of light is at 186,282 miles per second, which may be hard to envision for many people. So, we compare 186,282 miles with the perimeter of the Earth, which is 24,901 miles long. Going one round on the Earth takes 24,901 miles, and we divide 186,282 by 24,901:

$$186,282/24,901 = 7.48.$$

So going 186,292 miles is about as far as going around the Earth 7.5 times. By saying that light can travel around the Earth approximately 7.5 times in a second. We may relate better to how fast it is because we have a sense that the Earth is pretty large and hard to get around.

Here is another example. By late September 2020, the number of confirmed COVID-19 cases in the US surpassed 7 million. To get a sense of the size of 7 million people, we can compare 7 million people to the population of Minnesota (5.7 million) and the two Dakotas (ND: 0.8 million; SD: 0.9 million) combined. Depending on your geographic region, you can also compare to Arizona (7.2 million), Washington (7.6 million), or Los Angeles (4 million) and Chicago (2.7 million) combined.

*Using a Visual Perception*

Once there was a Northern Chinese king with an army of millions. As they were about to cross the Yangtze River (the longest and widest river of China) to conquer the South, someone asked the king how many people he brought. Instead of giving out the number, he famously said that his army was large enough *to block the entire Yangtze River if all the soldiers threw their horse whips in the river.* The king's conquest was not successful, but he was forever remembered for this quote. What he did here was converting a large number to a vivid visualization using a hypothetical assumption.

*Example: 200,000 People*

By September 2020, the number of COVID-19 related deaths in the US surpassed 200,000. How does one make sense of 200,000 people? Here is an idea. An average person's arm span is approximately 68 inches. If $200,000$ people held hands together to form a line, the line would be $200,000 \times 68 = 13,600,000$ inches long, or 215 miles. This is approximately the distance between Minneapolis and Fargo.

So, if all the people in the US who died from COVID were to hold hands one by one, they would form a line from Minneapolis to Fargo. For folks living elsewhere, some other possible references include: New York City to Washington DC (220 miles), Houston to Dallas (230 miles), Seattle to Portland (170 miles), etc.

### *Discussions*

1. Suppose you sometimes handle very large or tiny numbers at work. Find the scientific notations for each of the following numbers.

   a) 0.000000000000923

   b) 0.00000057

   c) 0.015

   d) 2.53

   e) 3152

   f) 18904564

   g) 50098003219428000

2. Find a very large number and try to make sense of it. Quote your resources, do your own calculation and make a catchy title or punchline.

   **Sample.** *How far is the Moon? You won't hear an explosion on the Moon until* 13 *days later. The distance from the Earth to the Moon is* $238,900$ *miles, and the speed of sound is* 767 *miles per hour. We know that sound cannot travel in space, but if it could, and if there were an explosion on the moon, it will take* $238,900/767 = 311.5$ *hours, or almost* 13 *days for the sound to reach the Earth. (Source of data: Wikipedia.)*

## *6.2 The Tower of Hanoi*

Reference: Wikipedia: The Tower of Hanoi

### *The Tower of Hanoi*

This is a puzzle invented by the French mathematician Édouard Lucas in 1883.

Legend tells of an Indian temple in Kashi Vishwanath. The temple contains a large room with three time-worn posts and 64 golden disks of different sizes. At the beginning of time, all 64 disks were stacked around Post #1, in the order of sizes with the largest at the bottom. Brahmin priests, acting out the command of an ancient prophecy, have been moving these disks according to the following rules:

1. Only one disk can be moved each time from one post to another, and
2. No disk can be placed above a smaller disk.

According to the legend, when the entire stack of 64 disks is moved to another post, say Post #2 or Post #3, the world will end. (Picture: An 8-disk version.)

How many steps does it take to move the entire stack of 64 disks to another post? That is an interesting math problem. Here is how mathematicians approach this question.

1. To move the top disk to another post, it takes 1 step.
2. To move the top 2 disks to another post, we need to move the top disk away (1 step), move disk #2 (1 step) to another post and move the top disk back on top of disk #2 (1 step). So, the total is $1 + 1 + 1 = 3$ steps.
3. To move the top 3 disks to another post, we need to move the top 2 disks away (3 steps), move disk #3 (1 step) to another post and move the top 2 disks back on top of disk #3 (3 steps). So, the total is $3 + 1 + 3 = 7$ steps.
4. To move the top 4 disks to another post, we need to move the top 3 disks away (7 steps), move disk #4 (1 step) to another post and move the top 3 disks back on top of disk #4 (7 steps). So, the total is $7 + 1 + 7 = 15$ steps.
5. To move the top 5 disks to another post, we need to move the top 4 disks away (15 steps), move disk #5 (1 step) to another post and

move the top $4$ disks back on top of disk #5 (15 steps). So, the total is $15 + 1 + 15 = 31$ steps. And so on.

Whenever we want to move a stack that is one disk more, the number of steps becomes $2$ times more and plus $1$. By following this pattern, one can write up the sequence of numbers: $1$, $3$, $7$, $15$, $31$, $63$, $127, \ldots$

Based on this observed pattern, moving the entire stack of $64$ disks will require $2^{64} - 1$ steps. (Details omitted.) Using a calculator, we find this number to be approximately $1.845 \times 10^{19}$. (Note: Most calculators would display "$1.845E19$".)

That's a lot of steps, but the priests also have a lot of time to work on this task (since the beginning of the Earth). Would they complete this task soon and bring the world to an end?

*When Would the World End?*

There are $60$ seconds in a minute, $60$ minutes in an hour, $24$ hours in a day, and $365.25$ days in an average year (considering leap years). So, there are $60 \times 60 \times 24 \times 365.25 = 31,557,600$ seconds in a year. Assume that the Brahmin priests move one disk per second, then they can make $31,557,600$ moves in a year. To end the world, it will require $(1.845 \times 10^{19})/31,557,600 = 5.846 \times 10^{11}$ years to move the entire stack of 64 to another post.

According to science, the Earth is approximately $4.5$ billion, or $4.5 \times 10^{9}$, years old. Even if the priests began their work as early as the Earth was born, they still need $5.846 \times 10^{11} - 4.5 \times 10^{9} = 5.801 \times 10^{11}$ years to complete this job.

So, you should feel safe to go to sleep tonight! The world will not end in your sleep tonight or anytime soon.

*Discussion*

Play the game of the tower of Hanoi at https://www.mathsisfun.com. What do you find out about this game? Share your thoughts, ideas and any comment.

## 6.3 To Infinity... And Beyond!

When we count things one by one, we use numbers like $1$, $2$, $3$, $4$, $5$, etc. These numbers are called *counting numbers*, or *natural numbers*. The counting numbers start with $1$ and increase by $1$ each time.

Notice that counting numbers get bigger and bigger, and the list of counting numbers never ends. In fact, *you can never find a counting number* *that is the largest of them all.* The reason is quite simple: If you can find the largest counting number, say $N$, then $N + 1$ is also a counting number and is larger than $N$. This contradicts the assumption that $N$ is the largest. In other words, you can keep counting and you will keep getting larger and larger numbers. We say that the counting numbers are infinite.

### *About Infinity*

Infinity is an idea used to describe being *greater than any given number*, no matter how large that number is. In mathematics it is denoted by the symbol $\infty$, like an 8 lying sideways.

$\infty$ is not a number. If it were a number, you can add 1 to it and create an even bigger number, contradicting the assumption that it is greater than any given number. Indeed the word infinity is used to describe *an unlimited increase that goes beyond all bounds.* For example, we say that the counting numbers go by 1, 2, 3, 4, 5, and *to infinity.* This does NOT mean that there is a place or number called infinity and the list stops there. Instead, it means the list keeps going on and the numbers keep increasing and beyond any given number you can name.

### *The Infinitesimal*

Since counting numbers can get bigger and bigger, their reciprocals can get smaller and smaller, which means closer and closer to zero. For example, $\frac{1}{1000}$ is small, but $\frac{1}{2000}$ is even smaller because we use a larger denominator. It is quite intuitive that for any small number, we can always get an even smaller numbers. This is, however, not very easy to prove even for students in the math major.

We use the *infinitesimal* to describe the dynamic process of getting infinitely close to 0, just like we use the infinity to describe the process of getting greater than any given number. For example, the reciprocals of counting numbers are $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$, to *infinitesimal.* (Note: We cannot say they are $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ to 0 because none of these numbers is 0.)

### *The Clock Numbers*

One reason counting numbers keep going and eventually to infinity is because every number is followed by a larger number that has not appeared before. Not all number systems are like that. For example, consider the numbers on the clock: $1, 2, 3, 4, \ldots, 12$. If we list the numbers by recording the number after each hour, we will have: $1,\ 2,\ 3,\ 4,\ \ldots 12,\ 1,\ 2,\ \ldots$, which will NOT go to infinity.

## *Too Many to Count*

In mathematics, a collection of objects is called a *set* and each object is called an *element* of the set. For example, if we have a basketball team, the team is a set and each player is an element. A set is called *finite* when you can count its elements one by one, and complete the counting in a definitive number of steps. For example, the alphabets are a finite set of 26 elements. The student body of Minneapolis College is a finite set of a few thousand students. The population of the United States is a finite set of hundreds of millions of people.

Some finite sets are very large and may be technically difficult to count, but in concept they are still finite (like all the people in the US). Also, the order of counting doesn't really matter, as long as each element is counted exactly once.

When it is impossible to completely count all the elements, we call the set *infinite*.

### *Endless Counting*

Consider all the even numbers, listed from small to large: 2, 4, 6, 8, 10, …, 98, 100, 102, …. This set is infinite because we can never stop counting its numbers, as the list goes on and on. But on the other hand, every even number will eventually appear on the list. For example, the number 100 is the 50th to be counted, and the number 284 is the 142th. This is because 2, 4, 6, 8, 10, …, 98, 100, 102, … is a complete list of all the even numbers.

When an infinite set can be recorded on an endless but complete list, we call this set *countably infinite*, meaning that it can be counted one by one. For example, all the odd numbers $\{1, 3, 5, 7, \ldots, 99, 101, \ldots\}$ are countably infinite. All the multiples of 3 $\{3, 6, 9, 12, \ldots, 27, 30, 33, \ldots\}$ are also countable infinite.

Our next question is: Can a set have so many elements that it can't completely recorded on an endless list? In other words, can a set be *uncountably infinite*?

### *The Uncountable Infinity*

Mathematicians have been fascinated by the idea of infinity since the beginning of mathematics. They discovered the countable infinity early on, but it took them until the late 19th century to discover a higher order of infinity, the uncountable kind. The first uncountable infinity was discovered by the great Germany mathematician *Georg Cantor* (1845–1918 AD), who demonstrated that, all the numbers between 0 and 1 are uncountably infinite. The following is a simplified version of his discovery and the proof.

**Proposition.** No list of numbers can include all the numbers between 0 and 1.

*Proof.* (This is a proof by contradiction.) Assume that there is such a list of numbers $a_1, a_2, a_3, \ldots$ that has all the numbers between 0 and 1. We can write the entire list of numbers in decimals:

$$a_1 = 0.d_{11}d_{12}d_{13}\ldots$$

$$a_2 = 0.d_{21}d_{22}d_{23}\ldots$$

$$a_3 = 0.d_{31}d_{32}d_{33}\ldots$$

$$\ldots$$

Now we make a new decimal number $0.d_1d_2d_3\ldots$ by:

- Choosing a different 1st decimal from $a_1$: $d_1 \neq d_{11}$.
- Choosing a different 2nd decimal from $a_2$: $d_2 \neq d_{22}$.
- Choosing a different 3rd decimal from $a_3$: $d_3 \neq d_{33}$. And so on.

This number is different from all the numbers on the list, so it is NOT on the list. However, it is a number between 0 and 1 and should be on the list. This is a contradiction. □

### *Summary*

1. **Finite vs infinite.** A set is *finite* if we can record all its elements on a list that has an end. If not, we call it *infinite.*
2. **Two levels of infinity.** A set is *countably infinite* if we can put all its elements on an endless list. If not, we call it *uncountably infinite.*

### *Discussions*

1. **Clock Numbers.** The clock numbers are the twelve numbers on an analog clock: 1, 2, 3, 4, …, 11, 12.

   a) Between two clock numbers, can we determine which is small and which is large? If yes, how do we do it? If no, why not?

   b) Some people think it's a good idea that we replace 12 with 0, which represents noon or midnight. What's your take on this?

   c) Can you do algebra with clock numbers, like $2 + 5$ or $11 - 7$?

2. **The Days of the Week.** Give each day of the week a number: $0$ for Sunday, $1$ for Monday, $\ldots$, and $6$ for Saturday. Then the list of the days of the week in chronological order is $0, 1, 2, 3, 4, 5, 6, 0, 1, \ldots$. This is a sequence that keeps going but NOT to infinity.

   Find another example of a sequence that keeps going but not to infinity.

3. **Talking About Infinity.** In your own words, answer the following.

   a) Your 4th grader nephew told you that "I just learned about the biggest number ever, and it's called INFINITY!" How would you respond?

   b) Someone asks you what the difference between a very large number and infinity is. What would you say?

   c) Explain countable infinity vs uncountable infinity.

## *6.4 Videos: Paradoxes About Infinity*

Watch the following documentaries.

1. The Infinite Hotel Paradox - Jeff Dekofsky

2. What is Zeno's Dichotomy Paradox? - Colm Kelleher

### *Discussions*

1. **Hilbert's Infinite Hotel.** What did you learn about Hilbert's Infinite Hotel? Is there anything that impresses you?

2. **Zeno's Paradox.** What did you learn about Zeno's paradox? Is there anything that impresses you?

3. **Video overview.** Name one thing that is interesting enough for a casual conversation. ("Hey folks, I just learned this from my math class…"). Do it in an interesting enough way for such a conversation and report back what your friends say.

*Chapter 7*

# *Statistics Part 1 Tricked by the Graphs*

**Disclaimer.** This chapter is derived mainly from the course materials created by Dr. Cindy Kaus of Metropolitan State University.

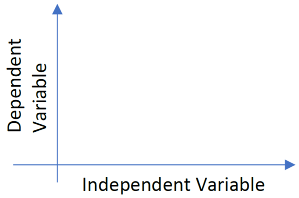
## *7.1 Graphing Basics*

A data is an information about an object. For example, assume you have three bowling balls: a 9-pound red ball, a 10-pound green ball and a 12-pound black ball. Their weight data (in pounds) can be recorded as $\{9,\ 10,\ 12\}$ and their color data can be recorded as $\{\text{red}, \text{green}, \text{black}\}$. (For clarity, we put the data inside the brackets.) We can also record both data using the notation (weight,color): $\{(9, \text{red}),\ (10, \text{green}),\ (12, \text{black})\}$.

*Variables*

Each category of the information recorded in the data is called a *variable*. With the bowling balls, there are two variables: the first one is the weight and the second one is the color of the ball. The weight variable is a number, so we call it a *numerical* (or *quantitative*) variable; the color variable is a quality and we call it a *qualitative* variable.

When the variables are numbers, we can observe when they are increasing or decreasing. For the data with two numerical variables, we can ask if and how the second variable changes with the first variable. For convenience we call the second variable *dependent variable* and the first one *independent variable*. On the plane we graph the independent variable in the horizontal direction and the dependent variable in the vertical direction.




*Basic Graphs*

For the following charts and graphs, we use the example of a bag of M&Ms with 16 Brown, 17 Red, 6 Yellow, 6 Green, 3 Blue and 2 Orange M&Ms.

**The frequency table** records how many times something appears. For the bag of M&Ms, the frequency table is

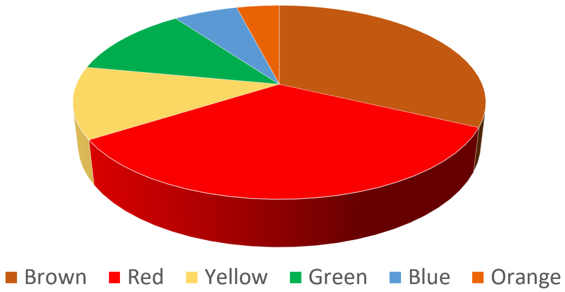
| Color | Frequency |
|---|---|
| Brown | 16 |
| Red | 17 |
| Yellow | 6 |
| Green | 6 |
| Blue | 3 |
| Orange | 2 |
| *Total* | 50 |

**The relative frequency table** records the percentage of the categories instead of the numerical counts of them. This is the *relative frequency*. With the M&Ms, there are 50 of them in total, so we divide the frequency of each category by 50. For example the brown M&Ms are $16 \div 50 = 32\%$.

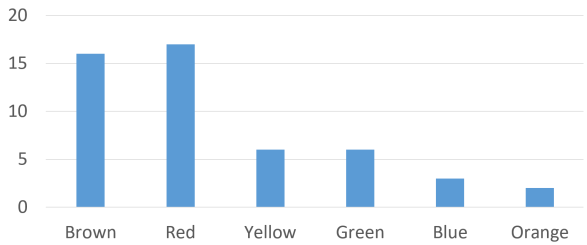
| Color | Relative frequency |
|---|---|
| Brown | 32% |
| Red | 34% |
| Yellow | 12% |
| Green | 12% |
| Blue | 6% |
| Orange | 4% |
| *Total* | 100% |

When calculating relative frequency, you sometimes need to round the answer. (For example, $3/122 = 0.0246901\ldots$, which you might round to 0.025 or 2.5%. Doing this to the entire table may result in the total percentages to be a little over or under 100%.

**The pie chart** is a graphic way of presenting the information on the relative frequency distribution. In a pie chart, each category is represented by a slice of the pie. The area of the slice is proportional to the relative frequency (percentage) of M&Ms in each category: The more percentage the larger slice of pie, and the total percentage is 100%.



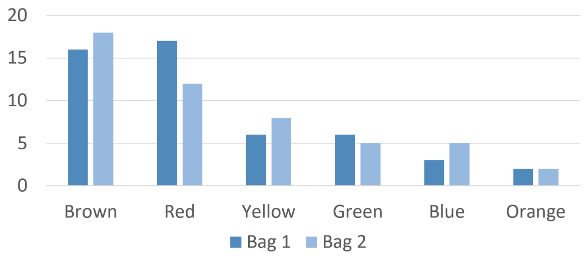

**Bar charts** is for qualitative categories and their frequencies or relative frequencies information. The categories go on the horizontal axis, and the frequency goes on the vertical axis. So, a bar chart is used to represent qualitative or categorical data. A bar chart of the M&Ms data:




**The combined bar charts.** Suppose we have two bags of M&Ms, then we can record both of them on the same bar chart by putting the bars side by side in different colors.

| Color | Bag 1 frequency | Bag 2 frequency |
|---|---|---|
| Brown | 16 | 18 |
| Red | 17 | 12 |
| Yellow | 6 | 8 |
| Green | 6 | 5 |
| Blue | 3 | 5 |
| Orange | 2 | 2 |
| *Total* | 50 | 50 |

The resulting bar chart:

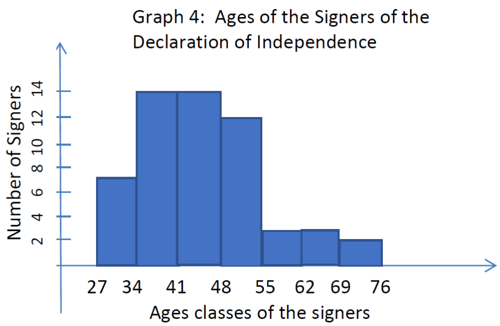



**Histograms.** When categories are not qualities (like colors) but a range of numbers, we use histograms. For example, the ages of the signers of the *Declaration of Independence* are shown in the table below.

| Age of signers | Frequency |
|---|---|
| 27–33 | 7 |
| 34–40 | 14 |
| 41–47 | 14 |
| 48–54 | 12 |
| 55–61 | 3 |
| 62–68 | 3 |
| 69–75 | 2 |

Note that the classes do not overlap, and each class runs right up against the next class. To make a histogram,

- Classes go on the horizontal axis (with the first number in each class denoted on the horizontal axis at the beginning of each bar).
- The frequency goes on the vertical axis The bars of the histogram are right next to each other. There are no gaps unless there is no data in that class.

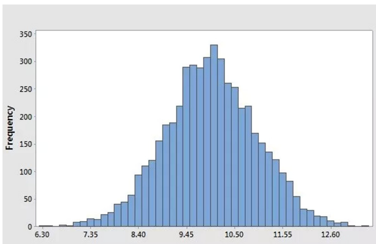


## *Special Shapes of Distribution*

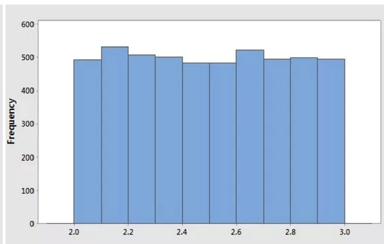
The histogram helps us visualize how things are distributed among the different ranges. There are some commonly seen patterns of distribution.

1. **Normally distributed**. This is also known as the bell curve.

2. **Uniformly distributed.** All the categories have (approximately) equal frequencies.



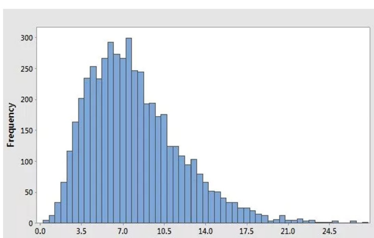

Normal Distribution



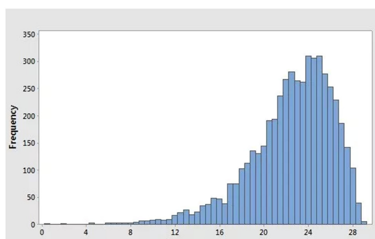

Uniform Distribution

3. **Skewed to the right.** A small part of the data has extremely high values, extending the graph to the right.

4. **Skewed to the left.** A small part of the data has extremely low values, extending the graph to the left.

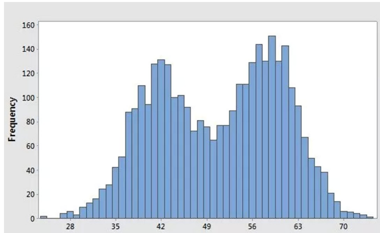



Skewed to the right




Skewed to the left

5. **Bimodal distribution.** There are two peaks for the distribution.




Bimodal

*Discussions*

1. **Finding Graphs in the News Media.** For the types of graphs introduced in the chapter, find one for each type in the news media. Tell the story and explain your finding.

2. **Special shapes of distributions.** For each of the 5 special shapes, find an example in the news media, tell the story and explain your finding.

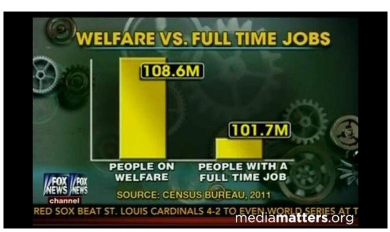
## 7.2 *Tricked by the Graphs*

A picture is worth a thousand word, as an old saying goes. But how about a misleading picture? We stay informed of our world by consuming data, graphs, and the information they present. When we look at a table or graph, we should ALWAYS ask ourselves the following questions:

- Where did this data come from? Is it from a reliable source?
- Is there a particular agenda of the group that is presenting this data? If there is, how might they be using the data to promote their agenda?
- What information is being left out of the presentation or included in the presentation that does not accurately portray the situation?

### *Graph 1: Welfare vs Full-Time Jobs*

The US Census Bureau published this graph in 2011:

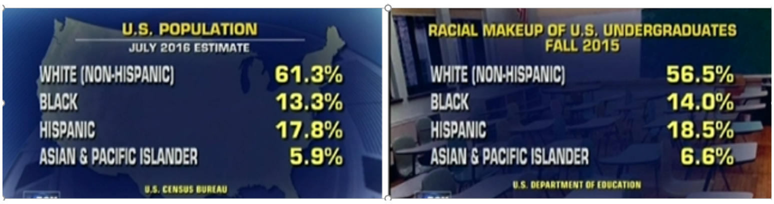


*Questions.*

1. Is there anything wrong with this graph?
2. Just for the context, what you think is the definition of welfare? What types of support are included in welfare and who is included as a welfare recipient?
3. How do you think the people represented in the graph below were counted?

### *Graph 2: Affirmative Action College Admissions*

A political group claims that affirmative action admissions policies at universities discriminate against white applicants. The following is its evidence.



| U.S. POPULATION JULY 2016 ESTIMATE | |
|---|---|
| WHITE (NON-HISPANIC) | 61.3% |
| BLACK | 13.3% |
| HISPANIC | 17.8% |
| ASIAN & PACIFIC ISLANDER | 5.9% |

U.S. CENSUS BUREAU

| RACIAL MAKEUP OF U.S. UNDERGRADUATES FALL 2015 | |
|---|---|
| WHITE (NON-HISPANIC) | 56.5% |
| BLACK | 14.0% |
| HISPANIC | 18.5% |
| ASIAN & PACIFIC ISLANDER | 6.6% |

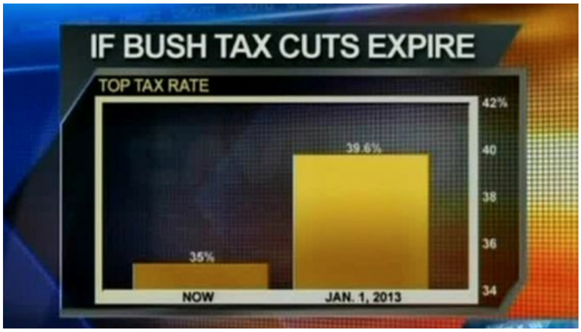
U.S. DEPARTMENT OF EDUCATION

*Questions*:

1. How do these tables support the political group's stance on affirmative action?
2. What information does the first table not consider?
3. To present an accurate comparison, what data should be presented in the first table?

### *Graph 3: George W. Bush Tax Cut*

This graph was used to support the reinstating of the Economic Growth and Tax Relief Reconciliation Act of 2001, tax cuts passed originally by George W. Bush in 2001. The graph shows the tax rate at 35% for the top income earners prior to 2013 and the tax rate at 39.6% for the top income earners in 2013, if the Bush tax cuts were to expire.

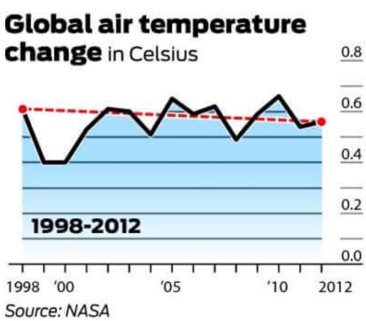


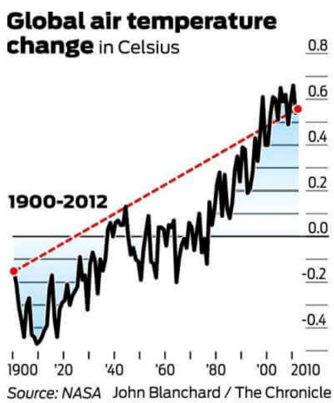

*Questions.*

1. Who is this graph trying to persuade?
2. How is this graph misleading? Look closely at the vertical axis.
3. Sketch a more accurate graph using the data that is given.

## *Graph 4: Climate Change*

Compare the following two graphs.



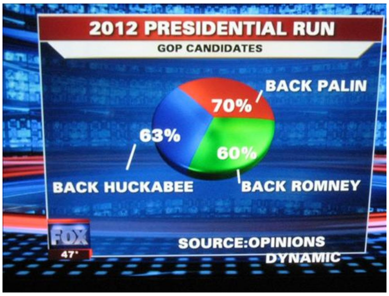

*Questions.*

1. Suppose I am a climate change denier. Which of the two graphs below would I use to support my view that the temperature of the earth is NOT rising? Why?
2. Suppose I am a climate scientist trying to get funding for my research on how to stop global warming. Which of the two graphs below would I use to support my view that the temperature of the earth is rising? Why?

3. Which graph is more accurate and why?

### *Graph 5: Not Just a Piece of Pie*

What is wrong with this graph?




## 7.3 *Census vs Sample*

Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data. (Merriam-Webster).

One important use of statistics is to collect data and present them. When we can obtain the information on ALL the involved objects, we have the so-called *census*. For example, the US census. When a census cannot be collected, statisticians collect a partial data and to try to understand the whole. This partial data is called a *sample*. For example, a poll for an upcoming election is a sample of the election, and the result of the poll is used to predict the outcome of the election.

### *Why Not Just Get the Census?*

One obvious reason is that it's easier and cheaper to study a sample than to examine the entire population. Also, some things are *rendered useless* after being surveyed or tested. For example, to find out about how long a light bulb would last, the light bulb must be kept on until it burns out. If a manufacturer wants to find out the life expectancy of its LED light bulbs, it should not test all of them.

### *Discussions*

1. **Census or Sample?** Present an interesting statistics information and identify if it is based on a census or a sample. Be sure to explain your reasoning and make the proper reference of your source of info.

2. **Statistics and You.** What is your impression and past experience with statistics? Have you used it before? Have you learn any statistics before college? Please share.

## 7.4 Video: The Best Stats You've Ever Seen

Watch this documentary: The Best Stats You've Ever Seen

### *Discussion*

1. Name two things you learned from this video. Please be specific.
2. Start a conversation with something mathematical and fun from this video. ("Hey folks, I just learned this from my math class...").

*Chapter 8*

# *Statistics Part 2 Describing Normal*

**Disclaimer.** This chapter is based on the course materials created by Dr. Cindy Kaus of Metropolitan State University.

## *8.1 The Center of the Data*

We hear conversations like these in daily life: An oil change *normally* takes an hour. It is *common* to spend $10 on a take-out meal. *Most people* need 10 to 14 days to recover from a viral infection.

In these sentences, there is a sense about being *normal* or *commonplace*, and we come to expect the data to generally be around a number, like one hour for an oil change. We call this number to expect the *center* of the data, and the commonplace behavior the *central tendency*. Depending on the context and the data, there are three ways to measure the center: the *mean* (aka the average), the *median* (the middle point of the data) and the *mode* (the most frequently occurring data).

**Example.** Employee Jack makes $60k a year. How well he does in his company should depend on how his $60k compares to the other employees. Let's consider three different Companies A, B and C, each having nine employees. For convenience, the salaries in thousands are arranged from low to high.

- Company A: 50, 60, 60, 60, 60, 60, 60, 60, 80.
- Company B: 10, 20, 20, 30, 30, 30, 40, 60, 80.
- Company C: 60, 70, 80, 80, 80, 90, 100, 100, 100.

If Jack is in Company A, he may feel \$60k is okay since almost everyone else earns \$60k. The "center" of the distribution is close to his salary.

If Jack is in Company B, he may feel \$60k is very good, since almost everyone else earns a salary under \$60k. The "center" of the distribution is below his salary.

If Jack is in Company C, he may be disappointed since everyone makes more than \$60k. The "center" of the distribution is above his salary.

### *The Mean*

The mean is the most common measure of center and the one that people are most familiar with. To find the mean of a data set, average them by adding up the values of the data and then dividing by the number of values in the data. For example, of the 5-point data set $\{35, 26, 49, 61, 73\}$, the mean is $\frac{35+26+49+61+73}{5} = 48.8$.

### *The Median*

The median is another common measure of center. It is the "middle" value in a data set. To find the median, you must arrange the data from lowest to highest value. If the data set has an odd number of values, the median is the exact middle. If the data set has an even number of values, the median is the average of the two middle values.

**Example.** Find the median for each of the data sets.

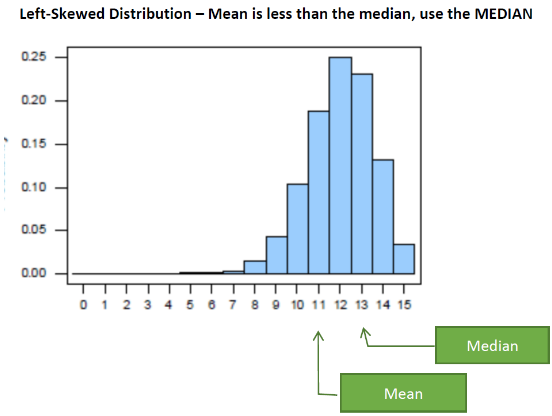
- 18, 11, 13, 2, 7, 7, 17, 20, 4
- 130, 120, 125, 139, 140, 122

*Solution.*

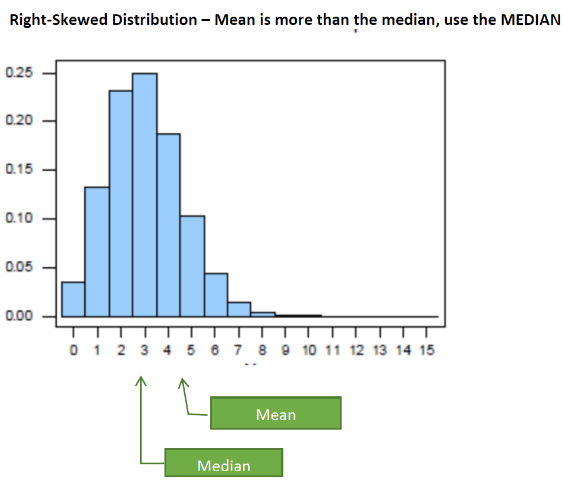
1. Ordering the numbers: 2, 4, 7, 7, 11, 13, 17, 18, 20. It is clear that the number in the middle is 11.
2. Ordering the numbers: 120, 122, 125, 130, 139, 140. There is not a middle number but a *middle space* between 125 and 130. So we average them: $(125 + 130)/2 = 127.5$ is the median.

**Skewed to the left/right.** The mean can be easily brought down by a tiny portion of uncharacteristically small data. This is the so-called *skewed to the left.* For example, you signed up for a class in week 2 and missed the first quiz, but you have been perfect for 4 weeks straight. So your quiz scores in the first 5 weeks are: $0\%, 100\%, 100\%, 100\%, 100\%$, which averages to $\frac{0\%+100\%+100\%+100\%+100\%}{5} = 80\%$. This does not reflect at all how well you do in this class.

Below is a example graph of a left-skewed data set.




Similarly, the mean can be easily brought up by a tiny portion of uncharacteristically large data. This is called *skewed to the right*. Below is a example graph of a right-skewed data set.




## *The Mode*

The mode of a data set is the value that appears the most often. If no value in your data set is repeated, then there is no mode for that data set. If more than one value appears the most often, both/all of them are modes. The mode is especially useful when there are a lot of repeated scores in the data, like the data set: {1, 1, 1, 1, 2, 2, 4, 400}. The mode captures the information that *the data is mostly 1's* instead of as a bunch of numbers averaging 51.5.

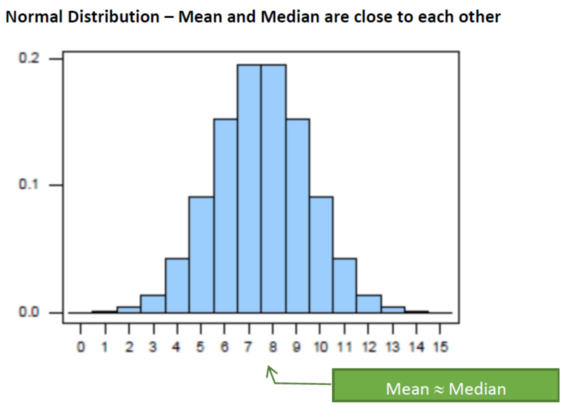
**Back to the example of Jack's salary.** Observe that:

1. Company A has mode 60, which appears 7 times.
2. Company B has mode 30, which appears 3 times.
3. Company C has modes 80 and 100, each appearing 3 times.

We can also calculate and find out that Company A has mean 61 and median 60, Company B has mean 36 and median 30, and Company C has mean 84 and median 80.

## *What Center to Use*

When you have a normal distribution, you can use either the mean or the median since they will be equal or close to equal.




When the distribution is skewed to the left or right (not symmetric), *the median is the preferred measure of center because the mean is affected by a tiny portion of uncharacteristically small or large data.*

## *Discussion*

1. **Central tendency in the news.** In this module, there are 3 ways of measuring central tendency: mean, median and mode. Find their examples in the news media. Be specific about the context, and highlight the part that is related to central tendency.
   ***Sample.*** *The* ***median*** *black family income in Minneapolis was \$36,000 in 2018, according to Census Bureau data. Though that figure compares*

*favorably with black families in many other U.S. metro areas, it is a far cry from the nearly $83,000 a typical white family in the city would earn. The $47,000 difference is one of the largest such gaps in the nation.*
*Source: The Washington Post*
*URL: ...*

2. **Living conditions in Minnesota.** Your friends from another state is considering moving their family to Minnesota, and they ask you how well families do in Minnesota. The following are three facts. Which one do you think best describes how well ordinary families do in 2018 in Minnesota? Explain your choice.

   a) The mean household income is $90,600. (US Census Bureau)

   b) The median household income is $68,411. (US Census Bureau)

   c) Approximately 12% are under the poverty line, $24,858. (MN Department of Health)

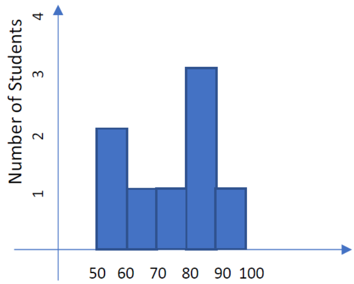
3. **Survey.** Have you learned or heard about the mean, the median and the mode in the past? Please share your prior experience or impression about this topic in statistics.

## *8.2 The Spread of the Data*

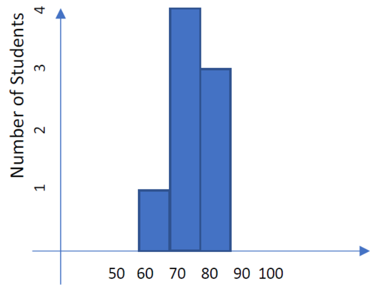
Mathematicians use the **standard deviation** to measure how much a data spreads out: the larger this number, the more spread out the data. In this section, we will follow the thought process that led mathematicians to the formula of standard deviation. For simplicity, we use the following example of three math classes and their test scores.

- Class A: 8 students, with scores 53, 55, 69, 75, 85, 85, 87, 91.
- Class B: 8 students, with scores 69, 71, 71, 72, 73, 80, 81, 83.
- Class C: 2000 students. Half of them have 74 points and the other half have 76 points.

Let's calculate the mean for each class: Class A is $(53+55+69+75+85+85+87+91)/8 = 75$, Class B is $(69+71+71+72+73+80+81+83)/8 = 75$, and Class C is $(74 \times 1000 + 76 \times 1000)/2000 = 75$. So, the three classes have the same mean. Between Classes A and B, B is more "close together" (see the graph), while Class C is the most together because all the data points are just 1 point away from the mean.




A




B

How do we measure this *togetherness* of the data?

### *Total Square Difference*

Math's first attempt to measure this togetherness is by squaring the difference between each data point and the mean (we call this a *square difference*) and adding up these square differences. The idea is that the closer the data is to the mean, the smaller this result will be. We calculate this total square difference for Classes A, B and C:

- Class A: $(53-75)^2+(55-75)^2+(69-75)^2+(75-75)^2+(85-75)^2+(85-75)^2+(87-75)^2+(91-75)^2=1520$.
- Class B: $(69-75)^2+(71-75)^2+(71-75)^2+(72-75)^2+(73-75)^2+(80-75)^2+(81-75)^2+(83-75)^2=205$.
- Class C: $1000\times(74-75)^2+1000\times(76-75)^2=2000$.

Class B is more together than Class A and it does have a smaller total square difference at 205. How about Class C? All its data points are just 1 point away from the mean 75, but its total square difference is the largest of the three because it has a lot more data points and their square differences add up to a large number. The total square difference fails to identify Class C as the most together.

### *The Variance*

Here is our next attempt. Instead of using the total square difference, we average it over the size of the data. This is the so-called *average square difference*, also known as the *variance* of the data. Again, we calculate the variance of the three classes.

- Class A: $1520/8=190$.

- Class B: $205/8 = 25.625$.
- Class C: $2000/2000 = 1$.

The new measurement now identifies Class C as the most *together* because it has the least variance. Class A is the most *spread out* and receives the greatest variance. So, the variance is a viable indicator of the togetherness: The smaller the variance, the more together the data.

### *The Standard Deviation*

The standard deviation is the square root of the variance. But why take the square root? Numbers can always be squared, but not so much for real world measurements with units. For example, $1\,\text{inch}$ is a length, but $1\,(\text{inch})^2$ is an area and something totally different. Also, $(76\,\text{points} - 75\,\text{points})^2 = 1(\text{points})^2$ has no meaning because "square point" has no meaning.

The variance becomes hard or impossible to comprehend with unit. One way to fix this is to take the square root and bring the unit back to the unit of the original data. We call this number the *standard deviation.* Back to the example,

- Class A has variance $190\,(\text{point})^2$, so its standard deviation is $\sqrt{190} = 13.78$ point.
- Class B has standard deviation $\sqrt{25.625} = 5.06$ points.
- Class C has standard deviation $\sqrt{1} = 1$ point.

The standard deviation is **the ultimate** viable indicator of the "togetherness", with an adequate unit and meaning.

*Discussion*

10 **Most Populated Counties in MN.** The population (in thousands) of the ten most populous counties of Minnesota are:

| County | Population |
|---|---|
| Hennepin | 1,266 |
| Ramsey | 550 |
| Dakota | 429 |
| Anoka | 357 |
| Washington | 262 |
| St. Louis | 199 |
| Stearns | 161 |
| Olmsted | 158 |
| Scott | 149 |
| Wright | 138 |

This data is small enough for using a calculator, but the easiest way is to use an Excel spreadsheet. You can easily find a tutorial on the internet, like here. Also, use (STDEV.P) or (STDEVP) in Excel.

1. Find the mean, the median, and the standard deviation of the population of these 10 counties.

2. Remove the top two counties and recalculate the mean, the median, and the standard deviation of the population of the remaining 8 counties.

3. Compare the two results above. What do you observe?

## *8.3 Percentiles*

In our daily life, we might hear things like: "*This baby girl's height is in the 95th percentile. She is so tall!* or "*Someone's salary is in the top 20 percent*". What do they mean?

For the baby girl, 95% of girls her age and born in an environment that supports optimal growth (as defined by the Center for Disease Control and the World Health Organization) are shorter than the baby. In other words, her height is *at the cutoff line* of the bottom 95% and the top 5%.

For the person whose salary is in the top 20%, he is *at or above the cutoff line* of the bottom 80% and the top 20%.

### *Quartiles*

Quartiles are the 25th-, 50th- and 75th-percentiles that can give us a glimpse of the spread of the data by looking at the data in four equal (or almost equal) parts.

### *The 68-95-99.7 Rule*

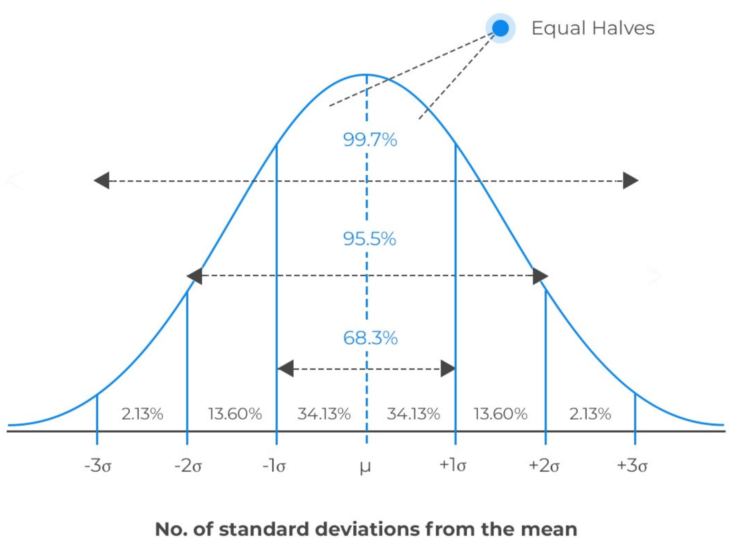
Source: Wikipedia. Edited for class use.

Photo credits: https://www.itl.nist.gov

The Normal Distribution is also known as the Bell curve. It is bell-shaped and symmetric with the center being the mean.

**The 68-95-99.7 Rule** is a shorthand used to remember the percentage of data within the mean in a normal distribution. To be specific,

- 68% of the data are within one standard deviation from the mean.
- 95% of the data are within two standard deviations from the mean.
- 99% of the data are within two standard deviations from the mean.




**An Example: IQ Test Scores.** An IQ test score is based on a norm group with an average score of 100 and a standard deviation of 15.

- One standard deviation is 15. So, 68% of people are within the range of $100 \pm 15$, which is between 85 and 115.
- Two standard deviations is 30. So, 95% of people are within the range of $100 \pm 30$, which is between 70 and 130.
- Three standard deviations is 45. So, 99.7% of people are within the range of $100 \pm 45$, which is between 55 and 145.

There are a lot of insightful information out of this. For example, 95% of people are between 70 points and 130 points, so only 5% of people are either below 70 points or above 130 points. So, 2.5% of people are below 70 points and 2.5% of people are above 130 points. So, someone with IQ 130 is at the top 2.5% of all people. In other words, 130 is the 97.5-percentile of the IQ Test.

### *Discussions*

1. **Percentiles in your life.** Find one usage of percentiles in news or social media. Does high percentile always mean better?

   ***Sample.*** *According to NFL DFS: DraftKings cheat sheet for Week 3, players considered to be a "Great" value are those in the 90th percentile of projected value picks for the week. "Good" plays are in the 75th percentile or better. "Average" plays span the 25th to 74th percentiles, while "Shaky" plays are in the 10th-24th percentiles. "Poor" picks encompass single-digit percentiles. Only players projected to score 4 points or more are displayed.*
   *Comment: Higher percentile means better players. Blah blah blah.*

2. **ACT scores.** The ACT math scores of all students are normally distributed with mean 20.5 with standard deviation 5.5. Answer the following questions.

   a) In the 68-95-99.7 Rule, how does the "68" range apply to this case?
   b) In the 68-95-99.7 Rule, how does the "95" range apply to this case?
   c) In the 68-95-99.7 Rule, how does the "99.7" range apply to this case?
   d) If someone scores 31.5, what percentile is this person?
   e) If someone scores 15, what percentile is this person?

## 8.4 Video: Own Your Body's Data

Watch this documentary: Own Your Body's Data

### Discussion

Name two things you learned from this video. Also create a conversation about the video.

*Chapter 9*

# *Statistics Part 3 The Linear Regression*

**Diclaimer.** This chapter is based on the course materials created by Dr. Cindy Kaus of Metropolitan State University.

## 9.1 The Relation Between Two Variables

Three kids, Jill, Hassan and Jose, went to an ice stand on a summer day. Jill ordered ice cream and fries. Hassan ordered chips and coke. Jose ordered ice cream. When their orders were ready, the server called: *Jill, ice cream. Jill, fries. Hassan, chips. Hassan, coke. Jose, ice cream.*

So, instead of saying: "Jill, here is the ice cream you ordered", and so on, the server communicated WHO ($x$) ordered WHAT ($y$) by calling out $x$ and then $y$. In mathematics, we would do exactly the same, but use the parentheses to keep a better record in writing. Let $(x, y)$ denote person $x$ and his/her food item $y$. Then the record of customers and their orders are $(\text{Jill}, \text{ice cream})$, $(\text{Jill}, \text{fries})$, $(\text{Hassan}, \text{chips})$, $(\text{Hassan}, \text{coke})$, $(\text{Jose}, \text{ice cream})$.

### *What is a Relation*

A relation is a connection between $x$ and $y$. In this example, the context of the relation is that *person $x$ ordered food item $y$*. When that happens, we record this information by writing $(x, y)$. Note that $(x, y)$ and $(y, x)$ are quite different. For example, Jill ordered fries but *fries didn't order Jill.*

On a data table (or a database), these orders can be organized like this:

| $x$: person | $y$: food item |
|---|---|
| Jill | ice cream |
| Jill | fries |
| Hassan | chips |
| Hassan | coke |
| Jose | ice cream |

When both variables $x$ and $y$ are quantitative (in other words, they are numbers), the database of their relation can also be used to observe the tendency of how they interact with each other.

### *The More $x$ the More $y$*

Ice cream cones at Dairy Queen (DQ) cost $1.50 each. If we call the number of ice cream cones $x$ and the total cost $y$, then the relation between $x$ and $y$ can be captured by this table:

| $x$ (number of cones) | $y$ (the total cost) |
|---|---|
| 1 | 1.50 |
| 2 | 3.00 |
| 3 | 4.50 |
| etc. | etc. |

It is clear that the more ice cream cones you order the more you need to pay. In other words, *the more $x$, the more $y$*, as $y$ increases with $x$. We call this an *increasing relation.*

Some commonly used increasing relations include:

1. **Proportional.** $y$ is $x$ times a certain number, like $y = 1.5x$ for the ice cream cone example.

2. **Linear with positive slope.** $y$ is a certain number adding a (positive) number of $x$. For example, you order $x$ pizzas at $20 each (tax included) and the restaurant charges $5 for delivery. The total cost $y = 5 + 20x$.

3. Other relations based on formulas like $y = 2^x$ (exponential), $y = x^2$ for positive $x$ (quadratic), etc.

Often times we see an increasing pattern, but not a simple formula to describe it, and that's okay. For example, at a summer fair, hot dogs are $6 each but there is a buy-one-get-one 50% off deal. If we call the number of hot dogs $x$ and the total cost $y$ then their relation should be:

| $x$ | $y$ |
|---|---|
| 1 | 6 |
| 2 | 9 |
| 3 | 15 |
| 4 | 18 |
| etc. | etc. |

### *The More $x$ the Less $y$*

Burgers are \$2 each and soda is \$1 each. You plan to spend \$10 on $x$ burgers and $y$ sodas.

| $x$ burgers | $y$ sodas |
|---|---|
| 0 | 10 |
| 1 | 8 |
| 2 | 6 |
| 3 | 4 |
| etc. | etc. |

You can see that, the more burgers you buy the less money you can spend on soda. In other words, *the more $x$, the less $y$*, or we say that $y$ decreases as $x$ increases. We call this a *decreasing relation.*

Some commonly used decreasing relations include:

1. **Inverse proportional.** $y$ is a (positive) number divided by $x$. For example $x$ numbers of kids try to evenly divide up 3 pizzas with each one getting $y$. Then $y = \frac{3}{x}$.

2. **Linear with negative slope.** $y$ is a certain number subtracting a (positive) number of $x$, like the burger and soda example, $y = 10 - 2x$.

Sometimes there is a decreasing pattern but not a simple formula to describe it. For example, you go to the fair with \$20, and hot dogs are \$6 each with a buy-one-get-one 50% off deal. If $x$ denotes the number of hot dogs and $y$ denotes the money you have left after paying for the hot dog. Then their relation should be:

| $x$ | $y$ |
|---|---|
| 1 | 14 |
| 2 | 11 |
| 3 | 5 |
| 4 | 2 |

### *Relations with Other or No Patterns*

Some relations have observable but not exactly increasing or decreasing patterns. For example, Minnesota's temperature by the month, $x$ is month and $y$ is the average low (in degrees Fahrenheit):

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
| $y$ | 4 | 12 | 24 | 36 | 49 | 56 |
| $x$ | 7 | 8 | 9 | 10 | 11 | 12 |
| $y$ | 63 | 61 | 51 | 39 | 25 | 11 |

This database shows that $y$ increases with $x$ in the first 7 months and decreases afterwards.

**Can't see a pattern?** Most relations in the real world either have hardly observable patterns or no pattern at all. For example, we survey 1000 adults with college degrees about their student debt ($x$) vs how many hamburgers they eat per week ($y$). The size of the database basically prohibits us from "eyeballing" the pattern! And there may not be a pattern at all. To understand the relation with a big size or weak pattern, we will need the help from statistics and computers.

### *Discussions*

1. Find an example for each of the following. Explain your answer.

    a) An increasing relation.

    b) A decreasing relation.

    c) A relation with some kind of pattern but NOT increasing or decreasing.

    d) A relation that has no clear pattern or any pattern at all.

2. Assume that we survey 1000 adults with college degrees about their student debt $x$ and how many hamburgers they eat per week $y$. Without statistical evidence or analysis, we will have to debate about whether $x$ and $y$ are related. Do you think someone's student debt ($x$) and the number of his/her weekly consumption of hamburgers ($y$) are related in any way? Why or why not?

## 9.2 *The Linear Regression*

You may have read reports with *correlation* like these. What do they mean?

- Study Suggests *correlation* between Heart Health and Optimism (US News and World Report, Jan. 2015)
- Study finds *correlation* between education, life expectancy (United Press International, July 2015)
- (There is a) close and significant *linear correlation* between chocolate consumption per capita and the number of Nobel laureates per 10 million persons. Conclusion: chocolate consumption enhances cognitive function. (New England Journal of Medicine, October 2012)

### *What Is Regression*

It is important to study the relations between two quantitative (numerical) variables (like years of education vs life expectancy) in the data. This relation is seldom clear or simple. Assume that the data is recorded as $(x, y)$ where $x$ and $y$ are variables. The goal is to describe the trend between $x$ and $y$ with a mathematical formula called *regression.* The regression usually won't be a perfect match of the data, but it is the best possible. The most commonly used regression is the linear function $y = ax + b$, and we call it the *linear regression* of the data.

Regression can be calculated by using computer programs. (We will learn how to do it on Excel in the next unit.) When using the computer for a linear regression, we receive two information:

1. The formula of the line, and
2. A number that measures how well the formula fits the data. This number is called *determination coefficient*, or $R^2$. It is between 0 and 1 and the higher it is the better the regression fits the data.

| $R^2$ | Strength of the regression |
|---|---|
| 0 to 0.3 | None or hardly any |
| 0.3 to 0.5 | Weak |
| 0.5 to 0.7 | Moderate |
| 0.7 to 1 | Strong |

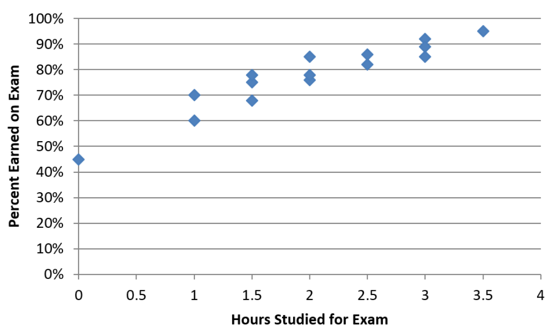
**Remark.** When variables are NOT numerical (like heart health and optimism), we will need to numerically measure it. (For example, ask people to rate their optimism from 0 to 10, etc.) Then we can do the regression of heart health vs optimism.

### *Study Time vs Grade Percentage*

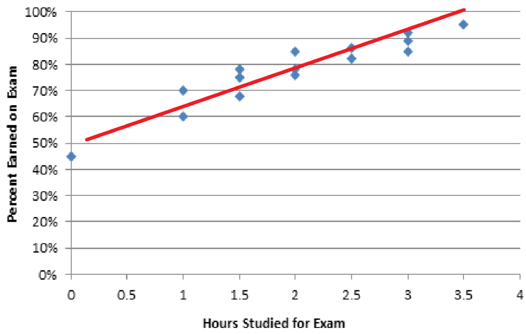
The following data was collected from 15 students in a math class after the first exam. The students were asked "How many hours did you study for the first exam?" and "What percent did you earn on the first exam?" The results are in the table below:

| Student | Hours studied | Percent earned |
|---|---|---|
| A | 2 | 76% |
| B | 0 | 45% |
| C | 3 | 92% |
| D | 2.5 | 86% |
| E | 1.5 | 75% |
| F | 1 | 70% |
| G | 1.5 | 78% |
| H | 3 | 89% |
| I | 3.5 | 95% |
| J | 3 | 85% |
| K | 2 | 78% |
| L | 2.5 | 82% |
| M | 1 | 60% |
| N | 1.5 | 68% |
| O | 2 | 85% |

We are interested in the relation between *the number of hours studied* and *the percent earned on the exam.* To get a picture of how these two variables are related, we graph the data by plotting $(\text{hour}, \text{percent})$ using $xy$-coordinates. This is called a *scatter plot.*



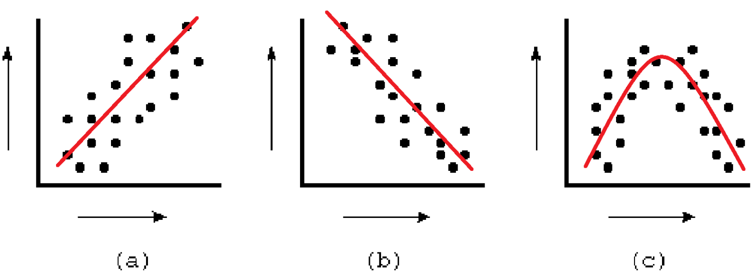

Do you see a pattern in the scatter plot? It seems the points fall in approximately a straight line.




The red line has a positive slope, so this is a linear relation with positive slope. Since all the data points lie closely to the red line, this is a *strong* correlation.

### *Different Type of Regressions*

There are different types of data trends, so there are different types of regressions. People normally use simple functions for regressions because they are easy to visualize and relate to. Below are some commonly used regressions: the positive linear regression, the negative linear regression, and the quadratic regression.




### *Back to Study Time vs Grade Percentage*

With the study of time $x$ vs grade percentage $y$, we can use Excel and find out:

1. The linear regression is $y = 0.1296x + 0.5168$, which helps us predict $y$ based on $x$. For example, if someone studies 1.2 hours a week, then the grade percentage is projected to be $y = 0.1296 \cdot 1.2 + 0.5168 = 0.67 = 67\%$. So, we predict that the person will earn 67% at the exam.

2. $R^2 = 0.9577$, which indicates a very strong correlation. This means we can be confident about the projection.

### *Finding Trends: Mathematics vs Statistics*

Consider the first six terms of the Fibonacci sequence $1, 2, 3, 5, 8, 13$. Can you predict the next term?

**The mathematician's approach** is to observe the patter that *every term is the sum of the previous two.* Therefore, we predict the next term to be $8 + 13 = 21$.

**The statistician's approach** is to do a regression of the six data points $(1, 1)$, $(2, 2)$, $(3, 3)$, $(4, 5)$, $(5, 8)$ and $(6, 13)$ and try to find out $(7, \text{WHAT?})$. In other words, the data points are:

| $x$: Order | $y$: Value | Data |
|---|---|---|
| 1 | 1 | (1,1) |
| 2 | 2 | (2,2) |
| 3 | 3 | (3,3) |
| 4 | 5 | (4,5) |
| 5 | 8 | (5,8) |
| 6 | 13 | (6,13) |

Using Excel, we find that $y = 2.286x - 2.667$ with $R^2 = 0.9$. With the next term, it is the 7th number ($x = 7$) and the statistician would predict $y = 2.286 \times 7 - 2.667 = 13.335$. You can see that the statistician totally missed the puzzle.

The linear regression is effective if the trend is linear or close to linear. But the linear regression assumes the simplest possible trend (the line) which doesn't suit other specially designed patterns. (Talking about hidden agenda, eh?)

### *Correlation or Causation?*

When two variables $x$ and $y$ are correlated, they follow a trend together. This may or may not mean that $x$ causes $y$ or $y$ causes $x$. For example, *the sales of umbrella* are correlated to *the number of rainy days*, and is indeed caused by it. However *the sales of umbrellas* and *the sales of raincoats* are correlated (as they both rise if the number of rainy days rises), but these two sales don't necessarily cause each other to increase or decrease.

Another famous example is that *the sales of ice cream and the number of murders in New York are positively correlated*, as both of them increase when temperatures are warmer. However, the sales of ice cream does not cause murders, and banning ice cream would do nothing to reduce the number of murders in New York City.

### *Discussions*

1. a) Find an example that $x$ and $y$ are highly correlated with $x$ being one of the causes of $y$.

   b) Find an example that $x$ and $y$ are highly correlated but they don't cause each other.

2. **Quantify!** To do regression, we need quantitative variables. Name three example of measuring a quality by the numbers. For example, intelligence is a quality that is commonly measured by the IQ score.

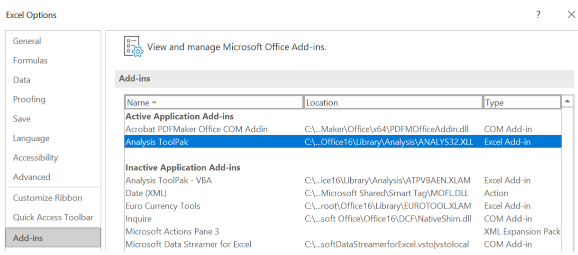
## 9.3 *Finding Regression with Excel*

Microsoft Excel spreadsheets can be used to analyze more data than what calculators normally handle.

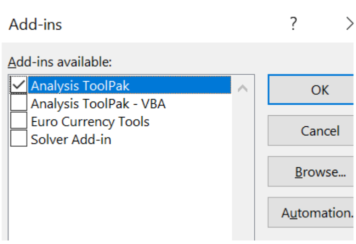
### *Setting Up Your Excel*

Follow these steps.

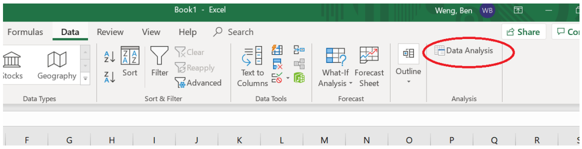
1. Go to File–> Options –> Add-ins and select Analysis ToolPak.




2. Choose to manage Excel Add-ins and select to activate Analysis ToolPak. Click OK.



3. Now go back to the main Excel window and choose Data. You should see a click button for Data Analysis.



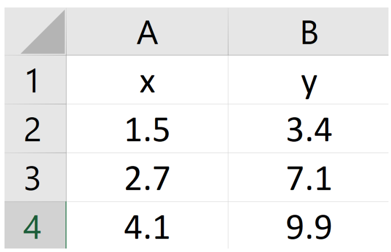

4. Your Excel is ready to go!

*Example*

Let's calculate the linear regression of a simple example:

| $x$ | $y$ |
|---|---|
| 1.5 | 3.4 |
| 2.7 | 7.1 |
| 4.1 | 9.9 |

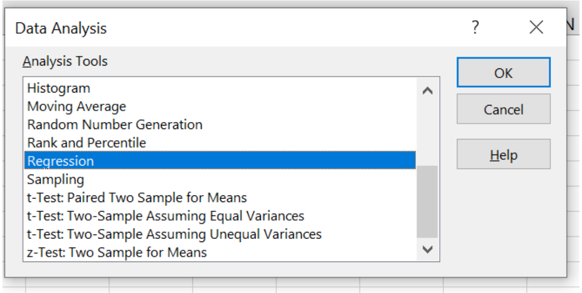
1. Enter the data on an Excel spreadsheet.



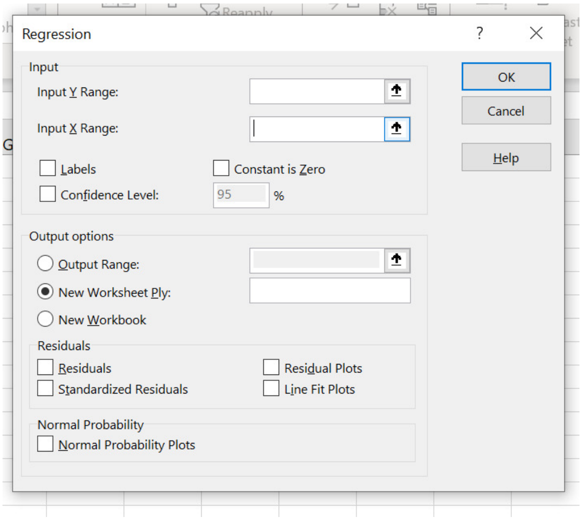
| | A | B |
|---|---|---|
| 1 | x | y |
| 2 | 1.5 | 3.4 |
| 3 | 2.7 | 7.1 |
| 4 | 4.1 | 9.9 |

Note that each space on the spreadsheet is called a cell. Every cell has a name (or a coordinate). For example, the first x value 1.5 is stored at A2.

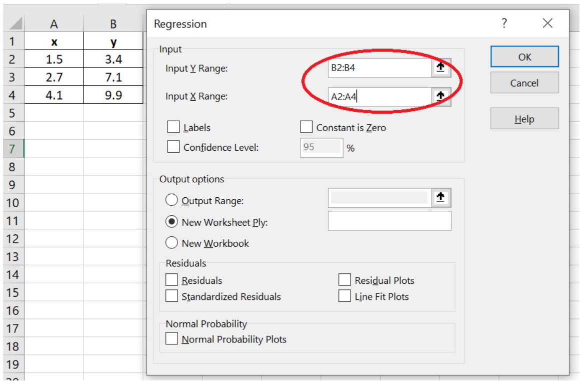
2. Choose Data and click on Data Analysis. On the roll-down menu, choose Regression.




3. We see this pop-up window:




4. Remember that each cell has a name. For the $Y$ range, we want the numbers 3.4, 7.1 and 9.9, so we enter **B2:B4**. Similarly, enter **A2:A4** for the $X$ range. Click OK.




5. The results of the calculation would be displayed on a separate sheet. The key information is highlighted in colored boxes:

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | SUMMARY OUTPUT | | | | | | | | | |
| 2 | | | | | | | | | | |
| 3 | *Regression Statistics* | | | | | | | | | |
| 4 | Multiple R | 0.992303 | | | | | | | | |
| 5 | R Square | 0.984665 | | | | | | | | |
| 6 | Adjusted F | 0.96933 | | | | | | | | |
| 7 | Standard I | 0.570985 | | | | | | | | |
| 8 | Observatio | 3 | | | | | | | | |
| 9 | | | | | | | | | | |
| 10 | ANOVA | | | | | | | | | |
| 11 | | *df* | *SS* | *MS* | *F* | *gnificance F* | | | | |
| 12 | Regressior | 1 | 20.93398 | 20.93398 | 64.21 | 0.079039 | | | | |
| 13 | Residual | 1 | 0.326024 | 0.326024 | | | | | | |
| 14 | Total | 2 | 21.26 | | | | | | | |
| 15 | | | | | | | | | | |
| 16 | | *Coefficients* | *andard Erro* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *ower 95.0%* | *pper 95.0%* | |
| 17 | Intercept | -0.07854 | 0.919534 | -0.08542 | 0.945754 | -11.7623 | 11.60525 | -11.7623 | 11.60525 | |
| 18 | X Variable | 2.48622 | 0.310269 | 8.013114 | 0.079039 | -1.45612 | 6.428561 | -1.45612 | 6.428561 | |
| 19 | | | | | | | | | | |

6. For the equation $y = mx + b$, the slope $m$ is the number in the blue box, $m = 2.48622$. The intercept $b$ is the number in the green box. $b = -0.07854$. So, the equation is $y = 2.48622x - 0.07854$.

7. The determination coefficient $R^2$ is the number in the red box, which is $0.984665$. This number is close to $1$, indicating a strong correlation.

## *Are Weight and Height Correlated in NFL?*

Here is a spreadsheet of the 2020 Vikings players: their player number, height, weight, age, and years in NFL. Using the technique we have learned about Excel, we can do a linear regression on weight $y$ vs height $x$. We find that the determination coefficient is $0.434521$, so the two variables are positively correlated, but not very strong. In other words, there is a trend that the taller (more $x$) a player is the heavier (more $y$) he would be, but this trend is not very strong.

*Discussions*

1. Each of the following sequence has a pattern. First find the pattern and predict the next number. Then use a regression to predict the next number.

    a) $1, 3, 6, 10, 15.$

    b) $1, -2, 3, -4, 5.$

    c) $1, 4, 9, 16, 25.$

    d) $1, -1, 1, -1, 1.$

2. **NFL Player Analysis.** Use the spreadsheet of Vikings players and the technique we have learned about Excel. Find a linear regression of the two given variables, and answer the following.

    a) This analysis is about $x$ vs $y$ about Vikings players.

    b) The linear regression is ....

    c) The determination coefficient is ....

    d) Based on the result, the strength of the correlation between $x$ and $y$ is ....

## 9.4 *Video: What Happens When Maths Goes Wrong?*

Watch the documentary: What Happens When Maths Goes Wrong? - with Matt Parker

*Discussions*

1. Name two things you learned from this video. Please be specific.

2. Name one thing that is interesting enough to start a casual conversation. ("Hey folks, I heard about this from my math class..."). Create such a conversation and report back what your friends say.

*Chapter 10*

# *Everybody Needs Continuity*

*Order, unity, and continuity are human inventions, just as truly as catalogs and encyclopedias. - Bertrand Russell*

What is continuity?

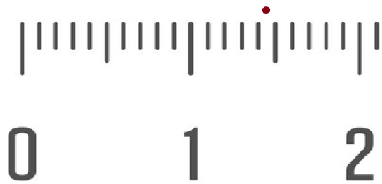
- You draw a line on a piece of paper. You call it *continuous* because it is *connected* and *without a break or gap.*
- Your job requires you to relocate from Minneapolis to Tokyo in 3 days. This drastic change causes a disruption of your life's *continuity.*
- You mix red and blue paints to make purple. To make the mixture slightly brighter, you add just a little more red paint because the red paint increases the brightness *continuously.*

Mathematicians use these intuitive ideas to define mathematical continuity.

## 10.1 The Continuum of Real Numbers

Lines appear to be connected in one piece, but that is a perception instead of the reality. If we draw a line with pencil, the black trace on the paper is the total of the carbon atoms, which are not scientifically attached to each other. A line on television screen is displayed by pixels lined up with (tiny) gaps between them.

Mathematics talks about lines and the other geometric objects in their abstract concepts; in that perspective we view lines as continuous. Greek mathematicians asked a question: *Can we identify every point on the line with a number?* For example, a ruler like this can identify points like $1.4$ and $1.5$, but not a point in between.




This question was natural for Greek mathematicians, who had figured out how to geometrically construct lengths that are fractions like $2/7$ and square roots like $\sqrt{2}$. The great Pythagoras thought the answer is yes because he believed that all numbers are ratios of whole numbers. Such numbers are called *rational numbers*, for example $6/13 = 6 : 13$, etc. In fact, Pythagoras made this belief the foundation of his theory of the physical and spiritual universe.

Unfortunately, Pythagoras was proven false three centuries later by another Greek mathematician Euclid, who showed that $\sqrt{2}$ is not a ratio between any two whole numbers. This is a mathematical milestone not only for the discovery itself, but also for the world premier of a logical method called *proof by contradiction*. This method would soon become an essential tool for math and logic.

Numbers that are not rational are called *irrational* numbers. (The words *rational* and *irrational* come from the word "ratio", not "ration".) Since Euclid, more and more irrational numbers have be identified, such as $\pi$, $e$ and the golden ratio $\phi$. With the rational numbers and all the newly discovered irrational numbers, can we establish a number system that can name every point on the line? That question remained unanswered for another two thousand years until the late 1800's, when mathematicians finally came up with a number system that is *continuous* like the line and can name all the points on it. This is the *real number system*. Through a fundamental theory called the*Axiom of Completeness of Real Numbers*, mathematicians establish the real number system, which has no "gap" and incorporates all the numbers we already have, from rational numbers to irrational ones like $\sqrt{2}$ and $\pi$.

## *Discussions*

1. **Bertrand Russell.** Who is Bertrand Russell? Find out about him and his influence in mathematics, philosophy and the modern society.

2. **Axioms.** Research on the internet about the axioms (also known as postulates) in mathematics. What are they? What role do they play

in Greek mathematics and all of mathematics? Report and comment on what you find out.

3. $\sqrt{2}$ **is irrational.** Legend goes that a follower of the School of Pythagoras discovered that $\sqrt{2}$ is irrational number before Euclid did three centuries later. Find out more about this story.

4. **The real numbers.** What do you know about the real numbers from your other math classes? Describe the real numbers in your own words.

5. **Pythagorean tuning in music.** Pythagoras was wrong about all numbers being rational, but his use of rational numbers influenced the human civilizations in many ways. Find out more about the Pythagorean tuning. What is it and what are its influences in music?

## 10.2 Continuous Functions

A function in mathematics is a process to generate an outcome $y$ from an input $x$. This type of processes is often described with formulas, like $y = 2x$, where the outcome is generated by doubling the input.

One thing people observe about *continuous changes* is that whenever $x$ changes gradually, so does $y$. So, whenever we change $x$ a little, $y$ would either remain the same or change only in a controllable amount. Functions behaving like this are called *continuous functions*. For example, $y = 2x$ is a continuous function; if we increase $x$ by $0.05$, $y$ will increase by $2 \times 0.05 = 0.1$.

How about another function, $y = 1000x$? If we increase $x$ by $0.1$, $y$ will increase by a lot more: $1000 \times 0.1 = 100$. But if we increase $x$ by a tiny bit at $0.0001$, $y$ would increase by $1000 \times 0.0001 = 0.1$. Compared to the first function, the outcome $y$ is a lot more *sensitive* to its input $x$.

### *Sensitivity*

With the two functions $y = 2x$ and $y = 1000x$, the two numbers $2$ and $1000$ are the measure of how fast $y$ changes with regard to $x$ ($2$ times and $1000$ times, respectively). These numbers are known in business and industry as the *sensitivity* of the functions, and in mathematics as the *first derivatives* of the functions. The study and the applications of the derivatives of functions are the major focus of the Calculus, an important advanced math tool for science and technology.

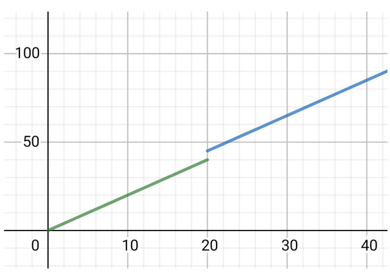
### *Discontinuous Functions*

Almost all the functions in your math classes are continuous, like the linear functions ($y = 2x - 5$), the quadratic functions ($y = x^2 + 4x - 3$), the exponential functions ($y = 2^x + 7$), the rational functions ($y = (x^3 + x - 5)/(x^2 + 1)$), etc. But we do see a lot of functions with discontinuities. Here are some examples.

**Bonus reward points.** You are in a consumer rewards program. You get 2 points for every dollar you spend, and 5 bonus points if you spend 20 dollars or more. Decimal points are allowed. Let $x$ be the money you spend, and $y$ be the points you receive. Before reaching 20 dollars, your reward points are twice the money you spend, so $y = 2x$. But at 20 dollars, your bonus points kick in. Let's demonstrate this in a data table:

| $x$ | $y$ |
|---|---|
| 19 | 38 |
| 19.9 | 39.8 |
| 19.99 | 39.98 |
| 20 | 40+**5** = 45 |
| ... | ... |

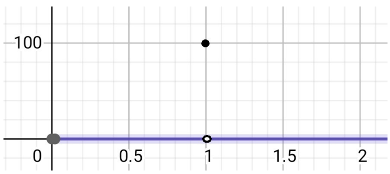
Therefore the function is NOT continuous when $x = 20$. We call this a *jump discontinuity* at $x = 20$. Here is its graph.




**An arcade game.** An old arcade game plays like this: A spotlight moves very fast in a circle, and you push a button to stop it. If you stop the light in front of the grand prize, you get the prize (like a SONY walkman), otherwise you get nothing. Let $x$ be the location of the light and $y$ be the result. When $x = 1$, the result $y = 100$ (because a walkman cost \$100), otherwise $y = 0$. The data table looks like this:

| $x$ | $y$ |
|---|---|
| 0.9 | 0 |
| 0.99 | 0 |
| 1 | **100** |
| 1.01 | 0 |
| 1.1 | 0 |
| ... | ... |

If we follow the tendency, when $x = 1$, $y$ should be 0, not 100. We call this a *removable discontinuity.* Here is its graph.




### *The Continuity in Life*

In the beginning we mentioned that relocating to Japan in 3 days is disruptive to life's continuity, but does it cause any discontinuity? From the mathematical perspective, the move causes your life to change a great deal in a short time, so it is sensitive to the time variable. However, it is not discontinuous: There may be a lot of change per day, but if you slow-play your life second-to-second, it is still a gradual and continuous process. An example of a discontinuity in life would be Captain Kirk and Commander Spock (in Star Trek) getting beamed down from the starship Enterprise to Earth. It happens in no time (or at light-speed), and no slow motion can make this process continuous.

### *Discussions*

1. **Example of continuity or discontinuity.** Given an example of a continuity or discontinuity from your experience. Describe what you mean by continuity or discontinuity in your example.

2. **Slope.** For the function $y = 1000x$, it is a linear function and 1000 is also called the *slope* of the function. From your past math classes, what do you know about the slope of a line?

3. **Star Trek.** In the example of Kirk and Spock, what type of discontinuity do they experience at the moment of beaming down to Earth? Explain your answer.

## 10.3 Meet Me in the Middle

The thermostat in your apartment is malfunctioning. Sometimes it makes the apartment too warm and other times too cold. Today, the temperature in your apartment was 60 degrees at 8 am and 85 degrees at 6 pm. Frustrated, your roommate said, "*This thing NEVER get the correct temperature!*" Is your roommate right about this?

Since 60 degrees is too low and 85 degrees too high, the ideal temperature for your roommate must be somewhere between 60 and 85. Because the real numbers are continuous, for the temperature to go from 60 to 85, it will hit all the number in between, including the ideal temperature for your roommate. (For example, you can't go from 60 to 85 continuously without hitting 70.) So, your roommate was wrong: *whatever the right temperature is, the apartment is at that temperature at least for a moment.*

### The Intermediate Value Theorem

The broken thermostat example demonstrates a mathematical principle called the Intermediate Value Theorem: If a continuous function has two different outcomes from two different inputs, then *there is always an intermediate input to generate a given intermediate outcome.* We don't always know what this intermediate input is, except that *it exists.* (A theorem is a mathematically proven proposition.)

#### Meeting in the Middle

The Intermediate Value Theorem describes a way of thinking we often apply unknowingly: Assume that you do something one way and the outcome is not enough, and you do it another way and the outcome is too much. Then somewhere in the middle you would for sure get it just right. For example, for your morning coffee, if one teaspoon of sugar is not sweet enough and two is too sweet, then the right amount of sugar is between one and two teaspoons.

The continuity of the functions and the continuity of real numbers play a crucial role in this principle. Without it, there may not be a way to create an outcome in the middle. For example, you wish to spend exactly $20 on burgers that are $3 each. 6 burgers cost less than $20, 7 burgers cost more than $20, but you can't buy $20/3 = 6\frac{2}{3}$ burgers to spend exactly $20.

The idea of finding something just right between "too much" and "not enough" can be used to solve equations. All you need to do is keep zooming in.

## *Searching for* $\sqrt{3}$

$\sqrt{3}$ is the number whose square is 3. Let's find it.

Consider a function $f(x) = x^2$. Our goal is to find the number that would generate the outcome 3. First, $f(1) = 1^2 = 1 < 3$ which is not enough, and $f(2) = 2^2 = 4 > 3$ which is too much. So somewhere between 1 and 2 we can get it just right. In other words,$1 < \sqrt{3} < 2$.

Next, we check all the one-decimal numbers between 1 and 2, from 1.1 to 1.9. We find that $f(1.7) < 3$ (not enough) and $f(1.8) > 3$ (too much). So $1.7 < \sqrt{3} < 1.8$. Then we check all the two-decimal numbers between 1.7 and 1.8, from 1.71 to 1.79, and repeat this process. This is time consuming, but not impossible with a calculator and some time. The more steps we take, the more decimal places we get for $\sqrt{3}$.

### *A Different Kind of Math*

The search for $\sqrt{3}$ uses an approach different from other techniques in algebra. Instead of using special tricks, we create a plan (algorithm) and execute it by performing numerical calculations. In mathematics, this approach is called a *numerical method*. The numerical methods became a dominant field of math in the twentieth century, with the rise of the computer technology. They continue to be a powerful tool for modern science and technology.

## *How Newton Solved* $\sqrt{3}$

Sir Isaac Newton (1642–1726 AD) was one of the greatest and most influential mathematicians and physicists of all time. By using Calculus, a new kind of math he invented, Newton created an efficient numerical method to solve equations. We will demonstrate how Newton's method can be used to find $\sqrt{3}$.

Newton's method begins with building a special function called the *iteration function*. Building the iteration function is beyond the scope of this class, but the iteration function for $\sqrt{3}$ is $f(x) = \frac{x}{2} + \frac{3}{2x}$. Next, we use the Intermediate Value Theorem to mark $\sqrt{3}$ between 1 and 2. Then we put in an estimated answer, say 1.5, into the function. The function produces an outcome that would be a better estimate of $\sqrt{3}$ than the input 1.5. To improve the result, keep feeding the outcomes back into the function for even better outcomes

1. Put 1.5 in the function: $f(1.5) = \frac{1.5}{2} + \frac{3}{2 \cdot 1.5} = 1.75$.
2. Put 1.75 in the function: $f(1.75) = \frac{1.75}{2} + \frac{3}{2 \cdot 1.75} = 1.73214$.

3. And so on.

How efficient is Newton's method? In two steps, we get an estimate of $1.73214$, which is accurate to three decimal places compared to the actual value at $1.73204\ldots$.

*Discussions*

1. Find an example in life that uses the principle of the Intermediate Value Theorem.

2. Find an example in life that cannot use the principle of the Intermediate Value Theorem.

3. **Isaac Newton.** Who is Isaac Newton? Find out about him and his influence in mathematics, physics, philosophy and to the world.

4. **Newton's method.** The following is a wanted number and its iteration function. Do three steps and compare your result with the actual value.

    a) $\sqrt{2}$. Use iteration function $\frac{x}{2} + \frac{1}{x}$, starting at 1.
    b) $\sqrt{5}$. Use iteration function $\frac{x}{2} + \frac{5}{2x}$, starting at 2.
    c) $\sqrt[3]{2}$. Use iteration function $\frac{2x}{3} + \frac{2}{3x^2}$, starting at 1.

## 10.4 Video: Can We Build a Brain

Watch this documentary: NOVA Wonders Can We Build a Brain?

*Discussions*

1. Name three things you learned from this video. Please be specific.

2. Name one thing that is interesting enough for a casual conversation. ("Hey folks, I just learned this from my math class…"). Do it in an interesting enough way for such a conversation and report back what your friends say.