



---

# CONTENTS

<b>1</b>	<b>Cognitive Biases</b>	<b>1</b>	<b>3</b>
1.1	Fundamental Attribution Error		3
1.2	Self-Serving Bias		4
1.3	In-group favoritism		4
1.4	The Bandwagon Effect and Groupthink		5
1.5	The Curse of Knowledge		5
1.6	False Consensus		6
1.7	The Spotlight Effect		7
1.8	The Dunning-Kruger Effect		8
1.9	Confirmation Bias		9
1.10	Survivorship bias		11
<b>2</b>	<b>Friction</b>		<b>13</b>
2.1	Static vs Kinetic Friction Coefficients		14
2.2	Skidding and Anti-Lock Braking Systems		16
<b>3</b>	<b>The Greek Alphabet</b>		<b>21</b>
<b>4</b>	<b>Basic Statistics</b>		<b>23</b>
4.1	Mean		23
4.2	Variance		24
4.3	Median		26
4.4	Histograms		27
4.5	Root-Mean-Squared		28

<b>5</b>	<b>Basic Statistics in Spreadsheets</b>	<b>31</b>
5.1	Your First Spreadsheet	31
5.2	Formatting	33
5.3	Comma-Separated Values	33
5.4	Statistics in Spreadsheets	34
5.5	Histogram	35
<b>A</b>	<b>Answers to Exercises</b>	<b>39</b>
	<b>Index</b>	<b>43</b>

# Cognitive Biases 1

In this section we're going to take a look at research findings about *cognitive biases*. These are universal quirks found in the human thought process. Cognitive biases aren't biases like racial biases. Everyone, regardless of nationality, race or gender is subject to these cognitive traps. You might be wondering, why do I need to learn about cognitive science in order to be an engineer? The most important tool we have as problem solvers is our own minds. We're going to be looking at ways that our minds can trip us up.

Our brains were designed over millions of years by the evolutionary process. The resulting mind is an amazing and powerful tool, however not flawless. The human brain has tendencies (or biases) that nudge us toward bad judgment and poor decisions.

When someone first gave you a hammer they handed it over with a warning, "don't hit your thumb!" No matter how careful you are with the hammer, at some point you'll hit your thumb. It's the same with cognitive biases. In the course of life all of us will fall prey to these cognitive biases. Knowing about them is the first step in protecting ourselves.

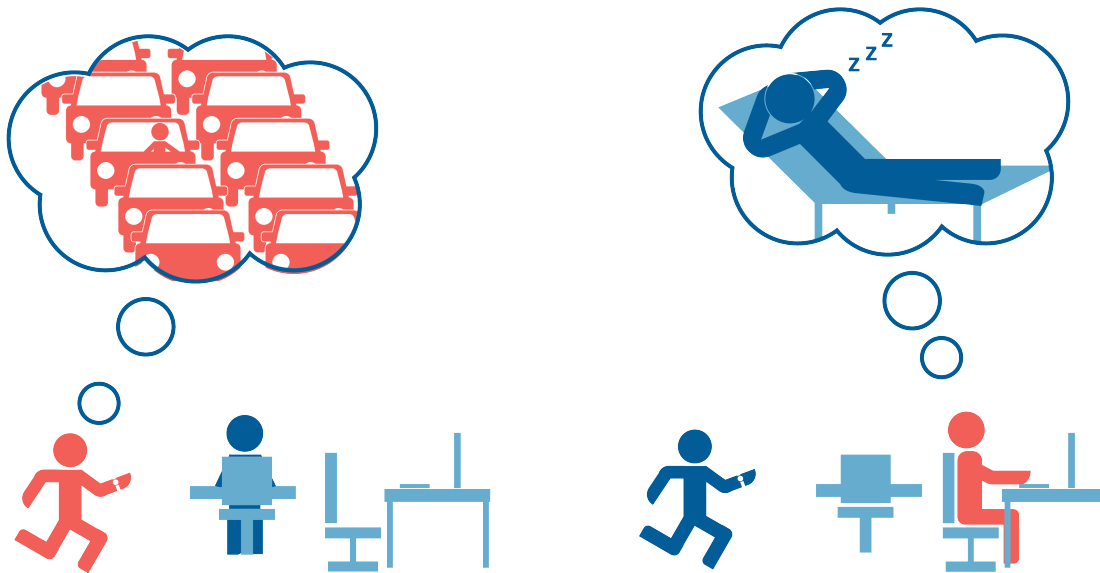
It would be irresponsible to teach you powerful ideas without also teaching you about the cognitive biases that follow them. There are about 50 that you should know about, but let's start with only a few.

## 1.1 Fundamental Attribution Error

You tend to attribute the mistakes of another person to their character, but attribute your own mistakes to the situation.

If someone asks you "Why were you late for work today?" You are likely to have an excuse, "I got stuck in a crazy traffic jam."

If you notice your coworker is late for work, you are likely to say "My coworker is lazy."



The solution? Cut people some slack. You probably don't know the whole story, so assume that their character is as strong as yours.

Or maybe you also need to hold yourself to a higher standard? Do you find yourself frequently rationalizing your bad judgment, lateness, or rudeness? This could be an opportunity for you to become a better person whose character is stronger regardless of the situation.

## 1.2 Self-Serving Bias

*Self-serving bias* is when you blame the situation for your failures, but attribute your successes to your strengths.

For example, when asked "Why did you lose the match?" you are likely to answer "The referee wasn't fair." When you are asked "Why did you win the match?" you are likely to answer "Because I have been training for weeks, and I was very focused."

This bias tends to make us feel better about ourselves, but it makes it difficult for us to be objective about our strengths and weaknesses.

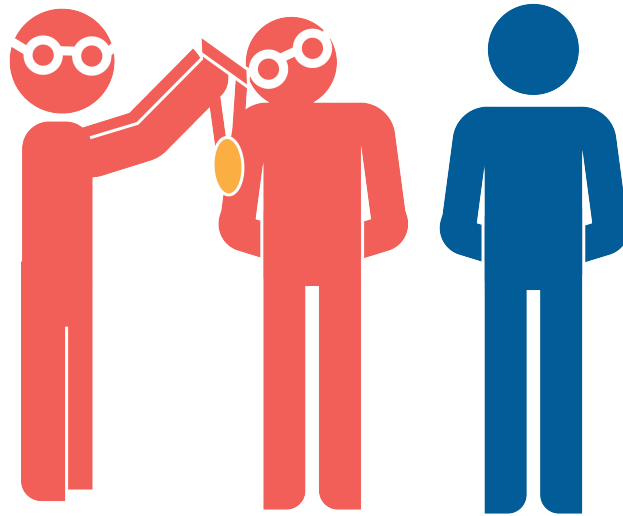
## 1.3 In-group favoritism

*In-group favoritism*: We tend to favor people who are in a group with us over people who are not in groups with us.

When asked “Who is the better goalie, Ted or John?”

If Ted is a Star Trek fan like you, you are likely to think he is also a good goalie.

As you might imagine, this unconscious tendency is the source of a lot of subtle discrimination based on race, gender, age, and religion.



## 1.4 The Bandwagon Effect and Groupthink

*The bandwagon effect* is our tendency to believe the same things that the people around us believe. This is how fads spread so quickly: one person buys in, and then the people they know have a strong tendency to buy in as well.

*Groupthink* is similar: To create harmony with the people around us, we go along with things we disagree with.

It takes a lot of perspective to recognize when those around us are wrong. And it takes even more courage to openly disagree with them.

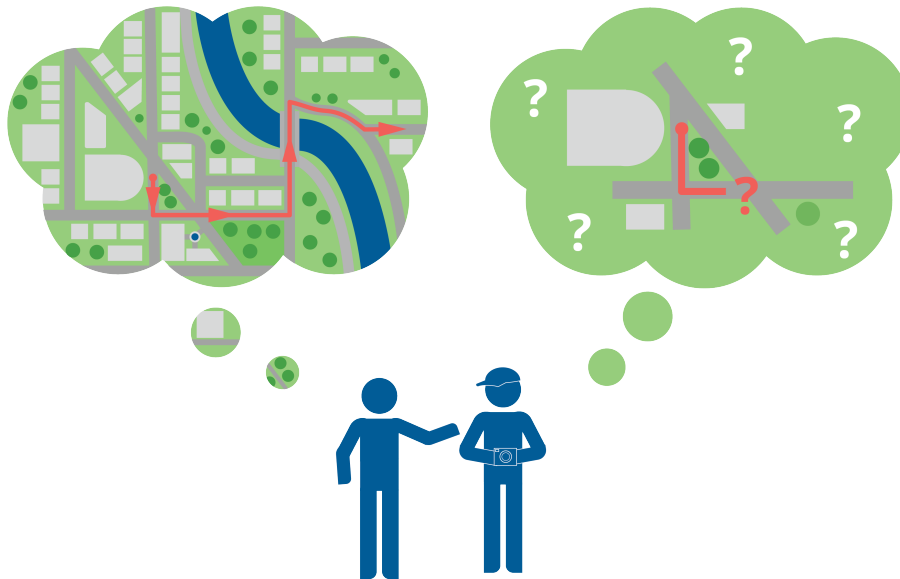
## 1.5 The Curse of Knowledge

Once you know something, you tend to assume everyone else knows it too.

This is why teaching is sometimes difficult: a teacher will assume that everyone in the

audience already knows the same things the teacher knows.

For example, imagine a local who has lived in a city for years giving directions to a tourist. The local has an in-depth understanding of the city, and gives overly quick and detailed instructions. The tourist politely smiles and nods, but stopped following after the local began listing unfamiliar street names and landmarks.



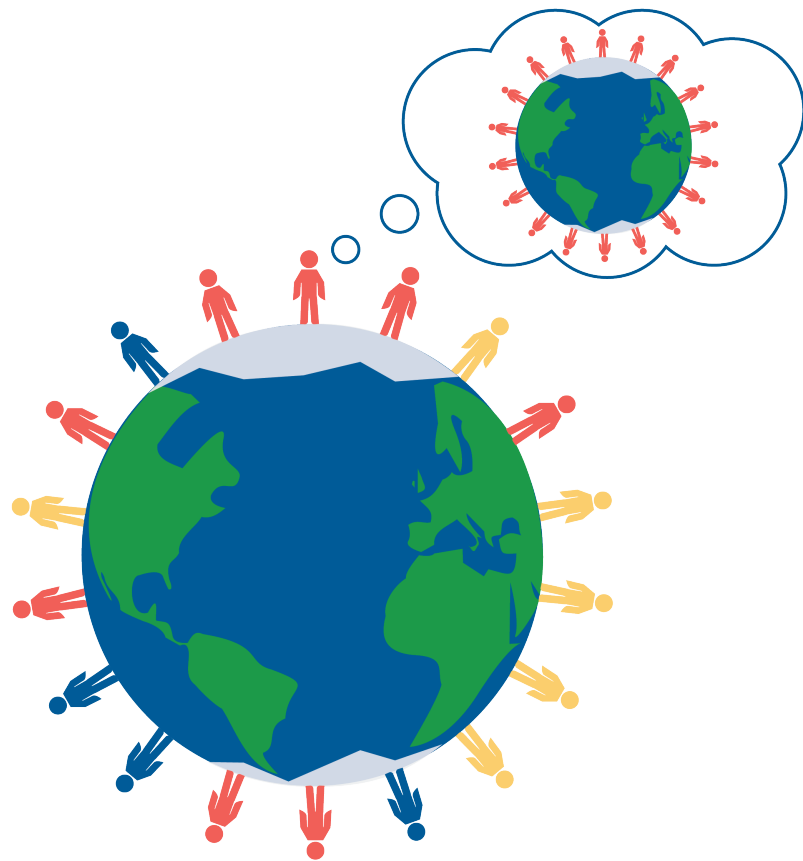
Also, when we learn that a friend doesn't know something that we know, we are often very surprised. This surprise can sometimes manifest as hurtful behavior.

When I find a gap in a friend's knowledge, I try to remind myself that the friend certainly knows many things that I don't. I also try to imagine how it would feel if they teased me for my ignorance.

## 1.6 False Consensus

We tend to believe that more people agree with us than is actually the case. For example, if you are a member of a particular religion, you tend to overestimate the percentage of people in the world who are members of that religion.

When people vote in elections, they are often surprised when their preferred candidate loses. "Everyone, and I mean EVERYONE, voted for Smith!" they yell. "There must have been a mistake in counting the votes."

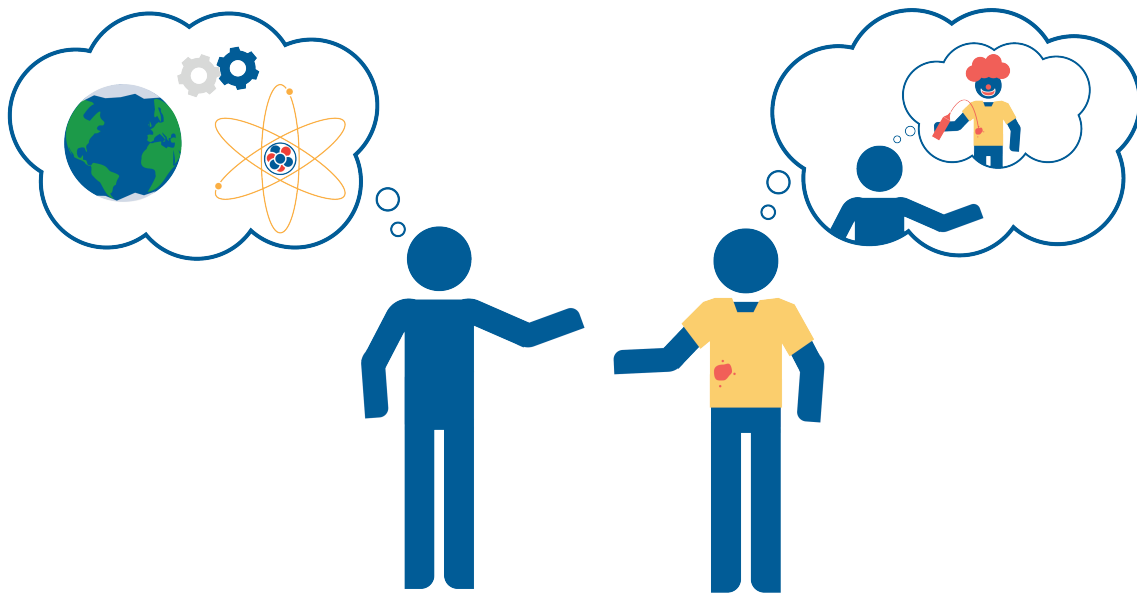


## 1.7 The Spotlight Effect

You tend to overestimate how much other people are paying attention to your behavior and appearance.

Think of six people that you talked to today. Can you even remember what shoes most of them were wearing? Do you care? Do you think any of them remember which shoes you wore today?

There is an old saying “You would worry a lot less about how people think of you, if you realized how rarely they do.”



## 1.8 The Dunning-Kruger Effect

The less you know, the more confident you are.

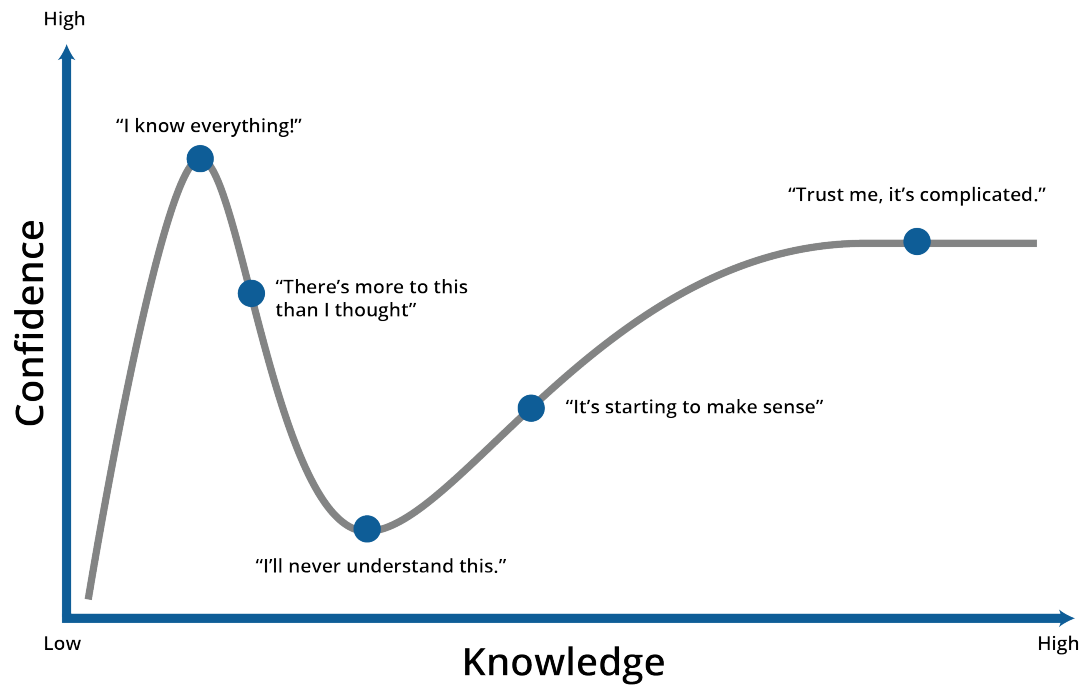
When a person doesn't know all the nuance and context in which a question is asked, the question seems simple. Thus the person tends to be confident in their answer. As they learn more about the complexity of the space in which the question lives, they often realize the answer is not nearly so obvious.

For example, a lot of people will confidently proclaim "Taxes are too high! We need to lower taxes." An economist who has studied government budgets, deficits, history, and monetary policy, might say something like "Maybe taxes *are* too high. Or maybe they are too low. Or maybe we are taxing the wrong things. It is a complex question."

When I am talking with people about a particular topic, I do my best to defer to the person in the conversation who I think has the most knowledge in the area. If I disagree with the person, I try to figure out why our opinions are different.

Similarly, you should assume that any opinion that is voiced, specifically, in an internet discussion is wildly over-simplified. If you really care about the subject, read a book by a respected expert. Yes, a whole book – there are few interesting topics that can be legitimately explained in less than 100 pages.



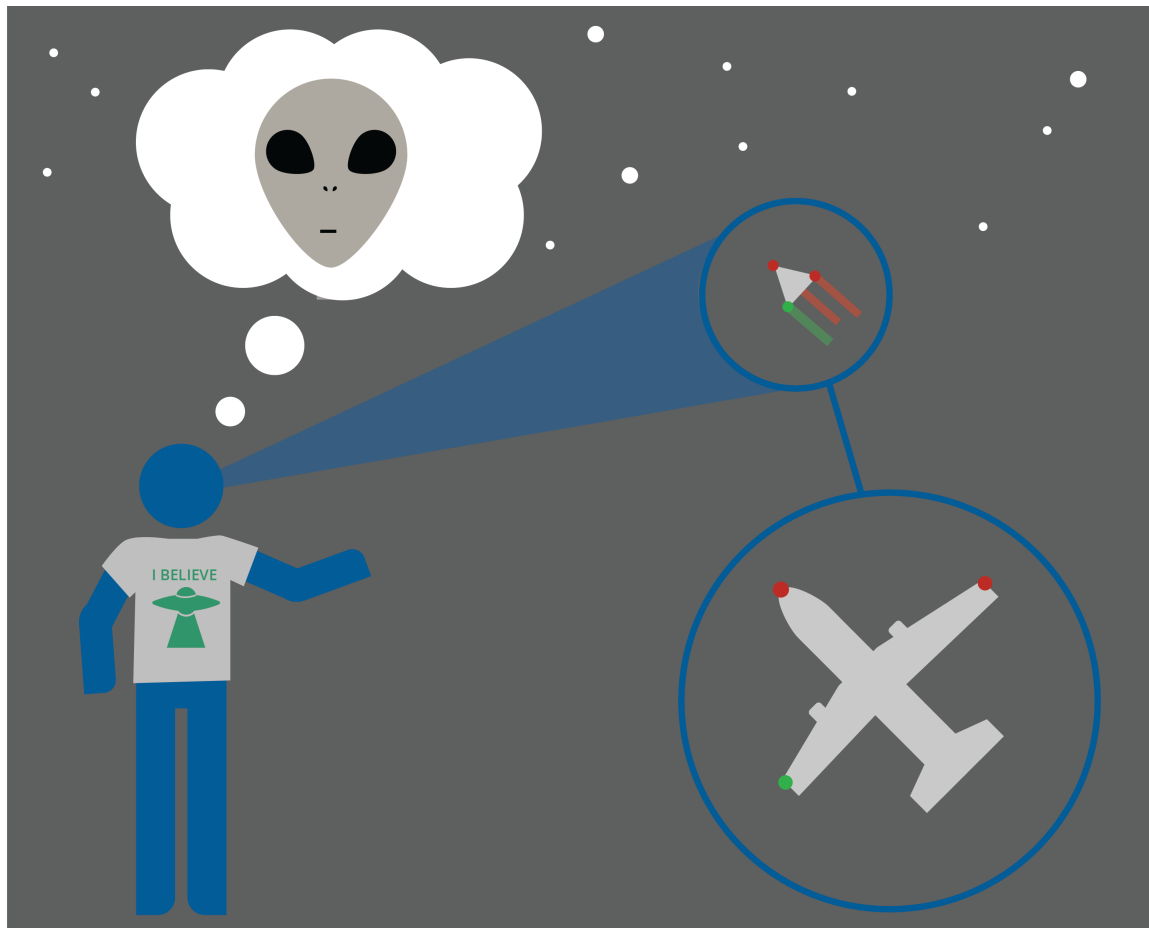


## 1.9 Confirmation Bias

You tend to find and remember information that supports beliefs you already have. You tend to avoid and dismiss information that contradicts your beliefs.

If you believe that intelligent creatures have visited from other planets, you will tend to look for data to support your beliefs. When you find data that shows that it is just too far for any creature to travel, you will try to find a reason why the data is incorrect.

Confirmation bias is one reason why people don't change their beliefs more often.



Confirmation bias wrecks many, many studies. The person doing the study often has a hypothesis that they believe and very much want to prove true. It is very tempting to discard data that doesn't support the hypothesis. Or maybe the person throws all the data away and experiments again and again until they get the result they want.

When you design an experiment, you must describe it explicitly before you start. You must tell someone: "If the hypothesis I love is incorrect, the results will look like this. If the hypothesis I love is correct, the results will look like that. And if the results look any other way, I have neither proved nor disproved the hypothesis."

Once the experiment is underway, you must not change the plan and you must not discard any data.

This is scientific integrity. You should demand it from yourself, and you should expect it from others.

Watch a TED Talk and Learn more about Confirmation Bias: What shapes our perceptions (and misperceptions) about science? In an eye-opening talk, meteorologist J. Marshall Shepherd explains how confirmation bias, the Dunning-Kruger effect and cognitive dis-

sonance impact what we think we know – and shares ideas for how we can replace them with something much more powerful: knowledge.

[https://www.ted.com/talks/j\\_marshall\\_shepherd\\_3\\_kinds\\_of\\_bias\\_that\\_shape\\_your\\_worldview](https://www.ted.com/talks/j_marshall_shepherd_3_kinds_of_bias_that_shape_your_worldview)



## 1.10 Survivorship bias

You will pay more attention to those that survived a process than those who failed.

After looking at a lot of old houses, you might say “In the 1880s, they built great houses.” However, you haven’t seen the houses that were built in the 1880s and didn’t survive. Which houses tended to survive for a long time? Only the great houses – you are basing your opinion on a very skewed sample.





## CHAPTER 2

---

# Friction

Imagine there is a large and heavy steel box resting in the middle of a large floor. Imagine you push it hard enough to get it moving. If you stop pushing, will it continue to glide gracefully across the floor?

Probably not. Unless the floor is very slippery for some reason, the box will come to a halt immediately after you stop pushing. We would say that it is stopped by the force of *friction*.

What's really happening? The kinetic energy of the box is being converted into heat between the bottom of the box and floor. As the bottom of the box and the floor get warmer, the speed of the box decreases.

The amount of friction is proportional to the force with which the box is pressing against the floor – so you should expect a box that is twice as heavy to experience twice as much frictional force.

That is, the frictional force is proportional to the normal force. (FIXME: picture here)

The amount of friction is also determined by the materials that are sliding against each other. For example, if the floor is ice, the frictional force will be less than if the floor is made of wood.

If you are pushing the box with a force of  $F$  and it is moving but neither accelerating nor decelerating, then the force you are applying is exactly balanced by the frictional force. If the box is pressing against the floor with a force of  $N$ , then we say the *coefficient of friction* between the steel box and the floor is given by

$$\mu = \frac{F}{N}$$

**Exercise 1**      **Bicycle Stopping**

*Working Space*

You are riding your bicycle at 11 meters per second when you slam on the brakes and lock up the wheels.

You weigh 55 kg.

When any piece of rubber is skidding across a dry road, the coefficient of friction will be about 0.7.

Answer the following questions:

- How much kinetic energy do you have when you engage the brakes?
- As you skid, how much frictional force is decelerating you?
- For how many meters will you slide?

*Answer on Page 39*

Notice that the force of friction is not determined by how much of the tire is touching the ground. The coefficient of friction of the two materials and the normal force all all you need to compute the friction.

## 2.1 Static vs Kinetic Friction Coefficients

Once again, imagine the box resting on the floor. As you start to push it, it will sit still until your force is greater than the force of friction. However, once it starts moving, the force of friction seems to be less.

Between two materials, there is actually 2 different friction coefficients:

- Kinetic friction coefficient: The coefficient you use once the box is sliding against the floor.

- **Static friction coefficient:** The coefficient you use to figure out how much force you need to get the box to start to move.

The kinetic friction coefficient is always less than the static friction coefficient:

- *Kinetic*,  $\mu_k$ : For a car skidding on a dry road, the friction coefficient is about 0.7.
- *Static*,  $\mu_s$ : When the car is parked with its brakes on, it has a friction coefficient of about 1.0.

## Exercise 2      Rocket Sled

Working Space

You are built a rocket sled with steel runners on a flat, level wooden floor. The sled weighs 50 kg and you weigh 55 kg.

Before you get on the sled, you push it around the floor some. You find that you can get it to move from a standstill if you push it with a force of 270 N. Once it is moving, you can keep it moving at the same speed using a force of 220 N.

**What are  $\mu_s$  and  $\mu_k$  of your sled's runners on your wooden floor?**

Now you get on the sled and gradually increase the thrust of the rocket mounted on the sled until it starts to move. Then you keep the thrust constant.

**How much force was the rocket exerting on you and the sled when it started to move?**

**How fast do you accelerate now that the sled is moving?**

Answer on Page 39

## 2.2 Skidding and Anti-Lock Braking Systems

When a car goes through a curve, the friction of the tire on the road is what changes the direction of the car's travel. Even though the wheel is turning, this is the static friction coefficient because the surface of the tire is not sliding across the road.

If you go into the curve too fast, the tire may not have enough friction to turn the car. In this case the car will start to slide sideways. Now the friction between the tire and road uses the kinetic coefficient. That is, you have significantly less friction than you had before you started to skid.

When you are driving a car, the force of friction that your tires create is your friend. It lets you steer, accelerate, and stop.

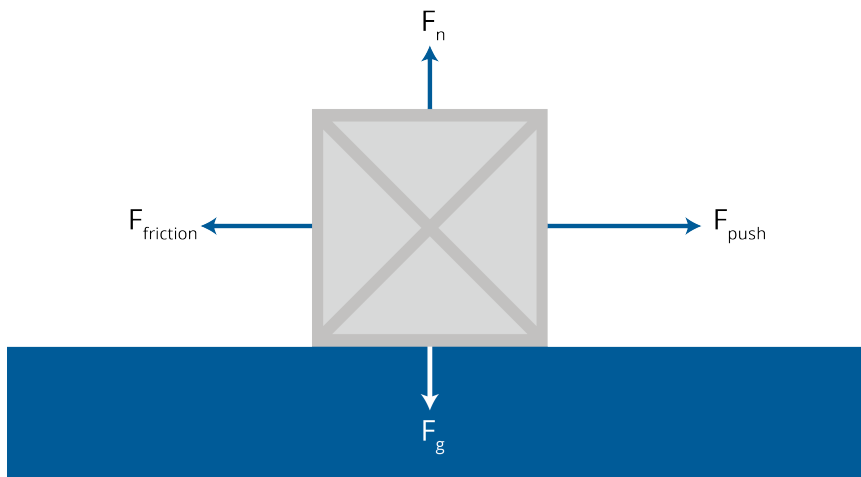
In older cars, if you would panicked and slammed on the brakes, you would probably lock up the wheels: they would stop turning suddenly. And the surface of the tire would begin to slide across the pavement. At that moment, two problems occurred:

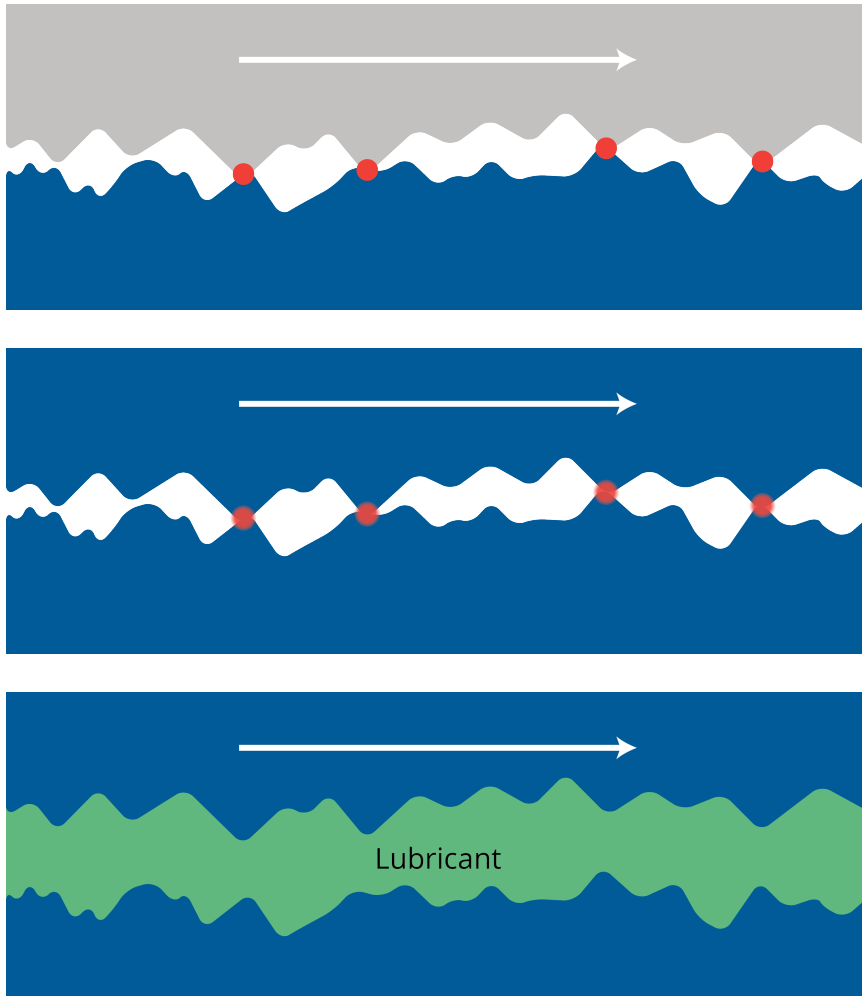
- You don't stop as quickly because now the friction between your tires and the road is based on the kinetic friction coefficient instead of the static friction coefficient.
- You can't steer the car. Steering only happens because the wheels are turning in a particular direction.

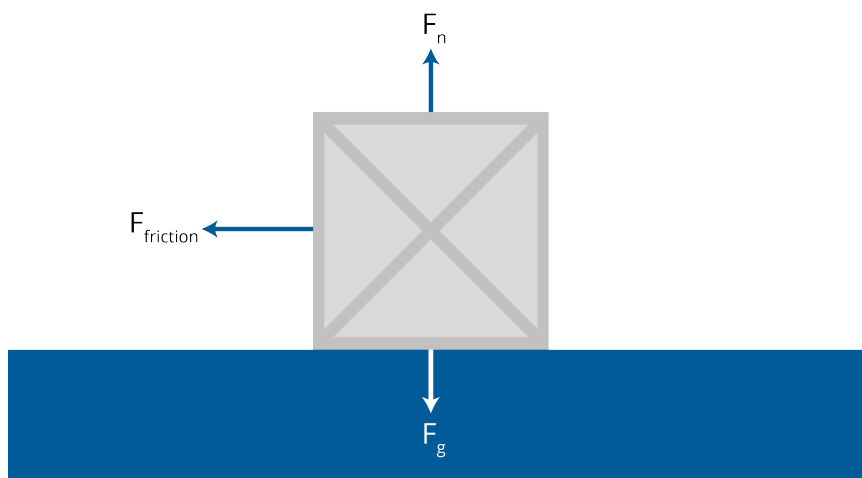
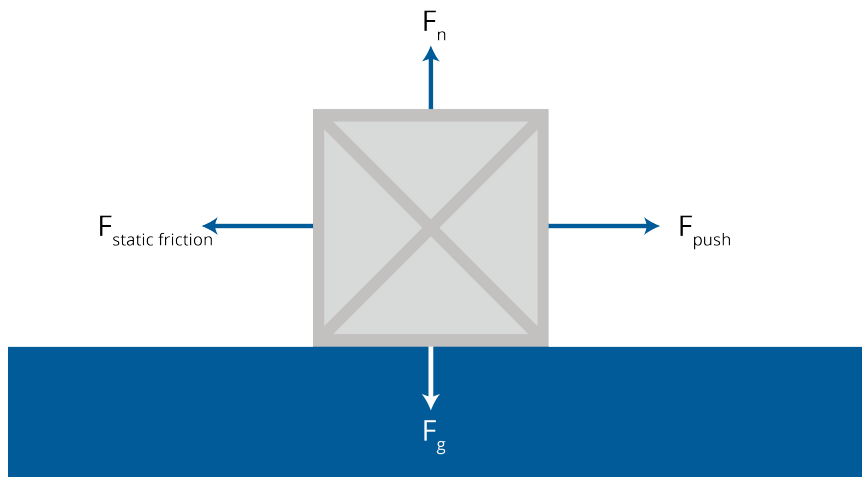
To prevent this problem, car companies developed the anti-lock brake system or ABS.

FIXME: More here.











# The Greek Alphabet

If you do anything involving math or physics, you will use a lot of Greek letters. Here is a table for your reference:

Capital	Lower	Pronounced	Capital	Lower	Pronounced
A	$\alpha$	Alpha	N	$\nu$	Nu
B	$\beta$	Beta	$\Xi$	$\xi$	Xi ("ku-ZY")
$\Gamma$	$\gamma$	Gamma	O	$\omicron$	Omicron
$\Delta$	$\delta$	Delta	$\Pi$	$\pi$	Pi
E	$\epsilon$	Epsilon	P	$\rho$	Rho
Z	$\zeta$	Zeta	$\Sigma$	$\sigma$	Sigma
H	$\eta$	Eta	T	$\tau$	Tau
$\Theta$	$\theta$	Theta	$\Upsilon$	$\upsilon$	Upsilon
I	$\iota$	Iota	$\Phi$	$\phi$	Phi
K	$\kappa$	Kappa	X	$\chi$	Chi ("Kai")
$\Lambda$	$\lambda$	Lambda	$\Psi$	$\psi$	Psi ("Sigh")
M	$\mu$	Mu	$\Omega$	$\omega$	Omega



# Basic Statistics

You live near a freeway, and someone asks you, “How fast do cars on that freeway drive?”

You say “Pretty fast.”

And they say, “Can you be more specific?”

And you point your radar gun at a car, and say “That one is going 32.131 meters per second.”

And they say, “I don’t want to know about that specific car. I want to know about all the cars.”

So, you spend the day beside the freeway measuring the speed of every car that goes by. And you get a list of a thousand numbers. Here is part of the list:

30.462 m/s	29.550 m/s	29.227 m/s
37.661 m/s	27.899 m/s	28.113 m/s
24.382 m/s	35.668 m/s	43.797 m/s
31.312 m/s	37.637 m/s	30.891 m/s

There are 12 numbers here. We say that there are 12 *samples*.

### 4.1 Mean

We often talk about the *average* of a set of samples, which is the same as the *mean*. To get the mean, sum up the samples and divide that number by the number of samples.

The numbers in that table sum to 388.599. If you divide that by 12, you find that the mean of those samples is 32.217 m/s.

We typically use the greek letter  $\mu$  (“mu”) to represent the mean.

#### Definition of Mean

If you have a set of samples  $x_1, x_2, \dots, x_n$ , the mean is:

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

This may be the first time you are seeing a summation ( $\sum$ ). The equation above is equivalent to:

$$\mu = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

### Exercise 3 Mean Grade

*Working Space*

Teachers often use the mean for grading. For example, if you took six quizzes in a class, your final grade might be the mean of the six scores. Find the mean of these six grades: 87, 91, 98, 65, 87, 100.

*Answer on Page 40*

If you tell your friend “I measured the speed of 1000 cars, and the mean is 31.71 m/s”, your friend will wonder “Are most of the speeds clustered around 31.71? Or are they all over the place and just happen to have a mean of 31.71?” To answer this question we use variance.

## 4.2 Variance

### Definition of Variance

If you have  $n$  samples  $x_1, x_2, \dots, x_n$  that have a mean of  $\mu$ , the *variance* is defined to be:

$$v = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$



That is, you figure out how far each sample is from the median, you square that, and then you take the mean of all those squared distances.

$x$	$x - \mu$	$(x - \mu)^2$
30.462	-1.755	3.079
29.550	-2.667	7.111
29.227	-2.990	8.938
37.661	5.444	29.642
27.899	-4.318	18.642
28.113	-4.104	16.839
24.382	-7.835	61.381
35.668	3.451	11.912
43.797	11.580	134.106
31.312	-0.905	0.818
37.637	5.420	29.381
30.891	-1.326	1.757
$\sum x = 386.599$ mean = 32.217		$\sum (x - \mu)^2 = 323.605$ variance = 26.967

Thus, the variance of the 12 samples is 26.967. The bigger the variances, the farther the samples are spread apart; the smaller the variances, the closer samples are clustered around the mean.

Notice that most of the data points deviate from the  $\mu$  by 1 to 5 m/s. Isn't it odd that the variance is a big number like 26.967? Remember that it represents the average of the squares. Sometimes, to get a better feel for how far the samples are from the mean, we use the square root of the variance, which is called *the standard deviation*.

The standard deviation of your 12 samples would be  $\sqrt{26.9677} = 5.193$  m/s.

The standard deviation is used to figure out a data point is an outlier. For example, if you are asked "That car that just sped past. Was it going freakishly fast?" You might respond, "No, it was within a standard deviation of the mean." or "Yes, it's speed was 2 standard deviations more than the mean. They will probably get a ticket."

A singular  $\mu$  usually represents the mean.  $\sigma$  usually represents the standard deviation. So  $\sigma^2$  represents the variance.

**Exercise 4      Variance of Grades**

*Working Space*

Now find the variance for your six grades.  
As a reminder, they were: 87, 91, 98, 65,  
87, 100.

What is your standard deviation?

*Answer on Page 40*

**4.3    Median**

Sometimes you want to know where the middle is. For example, you want to know the speed at which half the cars are going faster and half are going slower. To get the median, you sort your samples from smallest to largest. If you have an odd number of samples, the one in the middle is the median. If you have an even number of samples, we take the mean of the two numbers in the middle.

In our example, you would sort your numbers and find the two in the middle:

24.382	
27.899	
28.113	
29.227	
29.550	
<hr style="width: 100%;"/>	
<b>30.462</b>	
<b>30.891</b>	
<hr style="width: 100%;"/>	
31.312	
35.668	
37.637	
37.661	
43.797	

You take the mean of the two middle numbers:  $(30.462 + 30.891)/2 = 30.692$ . The median speed would be 30.692 m/s.

Medians are often used when a small number of outliers majorly skew the mean. For example, income statistics usually use the median income because a few hundred billionares

raise the mean a lot.

### Exercise 5 Median Grade

Find the median of your six grades: 87, 91, 98, 65, 87, 100.

*Working Space*

*Answer on Page 40*

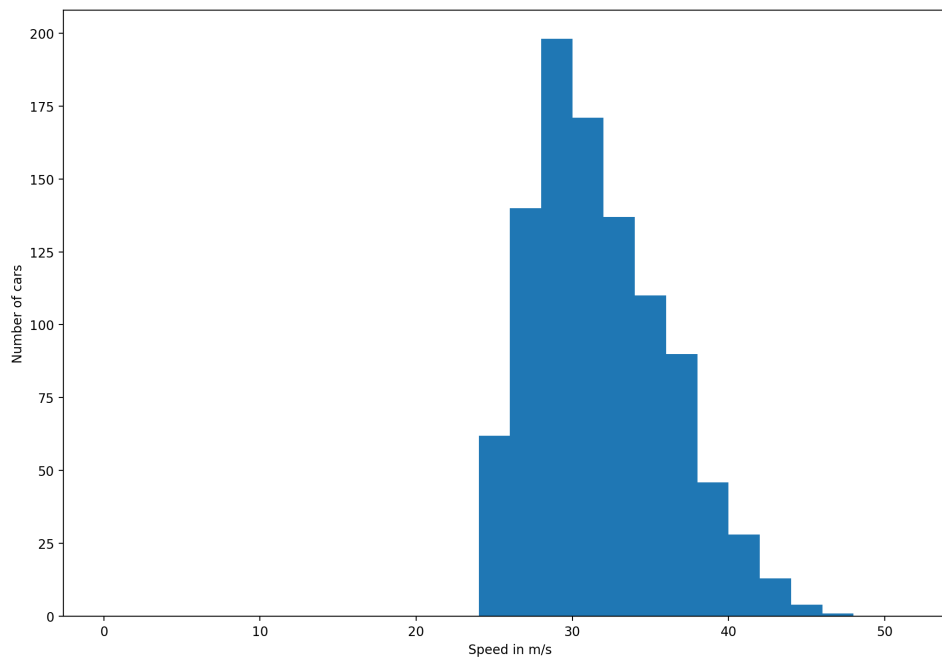
## 4.4 Histograms

A histogram is a bar chart that shows how many samples are in each group. In our example, we group cars by speed. Maybe we count the number of cars going between 30 and 32 m/s. And then we count the cars going between 32 and 34 m/s. And then we make a bar chart from that data.

Your 1000 cars would break up into these groups:

0 - 2 m/s	0 cars
2 - 4 m/s	0 cars
4 - 6 m/s	0 cars
...	...
20 - 22 m/s	0 cars
22 - 24 m/s	0 cars
24 - 26 m/s	65 cars
26 - 28 m/s	160 cars
28 - 30 m/s	175 cars
30 - 32 m/s	168 cars
32 - 34 m/s	150 cars
34 - 36 m/s	114 cars
36 - 38 m/s	79 cars
38 - 40 m/s	52 cars
40 - 42 m/s	20 cars
42 - 44 m/s	12 cars
44 - 46 m/s	4 cars
46 - 48 m/s	1 cars
48 - 50 m/s	0 cars

Now we make a bar chart from that:



Often a histogram will tell the story of the data. Here, you can see that no one is going less than 24 m/s, but a lot of people travel at 30 m/s. There are a few people who travel over 40 m/s, but there are also a couple of people who drive a lot faster than anyone else.

## 4.5 Root-Mean-Squared

Scientists have a mean-like statistic that they love. It is named quadratic mean, but most just calls it Root-Mean-Squared or RMS.

### Definition of RMS

If you have a list of numbers  $x_1, x_2, \dots, x_n$ , their RMS is

$$\sqrt{\frac{1}{n} (x_1^2 + x_2^2 + \dots + x_n^2)}$$

You are taking the square root of the mean of squares of the samples, thus the name Root-Mean-Squared.

Using your 12 samples:

x	x <sup>2</sup>
30.462	927.933
29.550	873.203
29.227	854.218
37.661	1418.351
27.899	778.354
28.113	790.341
24.382	594.482
35.668	1272.206
43.797	1918.177
31.312	980.441
37.637	1416.544
30.891	954.254
Mean of x <sup>2</sup>	1064.875
RMS	32.632

Why is RMS useful? Let's say that all cars had the same mass  $m$ , and you need to know what the average kinetic energy per car is. If you know the RMS of the speeds of the cars is  $v_{\text{rms}}$ , the average kinetic energy for each car is

$$k = \frac{1}{2}mv_{\text{rms}}^2$$

(You don't believe me? Let's prove it. Substitute in the RMS:

$$k = \frac{1}{2}m\sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2)}^2$$

The square root and the square cancel each other out:

$$k = \frac{1}{2}m\frac{1}{n}(x_1^2 + x_2^2 + \dots + x_n^2)$$

Use the distributive property:

$$k = \frac{1}{n}\left(\frac{1}{2}mx_1^2 + \frac{1}{2}mx_2^2 + \dots + \frac{1}{2}mx_n^2\right)$$

That is all the kinetic energy divided by the number of cars, which is the mean kinetic energy per car. Quod erat demonstrandum! (That is a Latin phrase that means "which is what I was trying to demonstrate". You will sometimes see "QED" at the end of a long mathematic proof.)

Now you are ready for the punchline: kinetic energy and heat are the same thing. Instead of cars, heat is the kinetic energy of molecules moving around. More on this soon.

Video: Mean, Median, Mode: <https://www.youtube.com/watch?v=5C9LBF3b65s>

# Basic Statistics in Spreadsheets

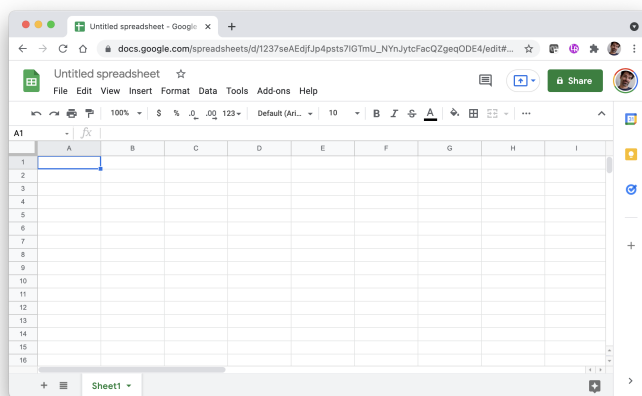
When you completed the problems in the last section, you probably noticed how long it took to compute statistics like the mean, the median, and variance by hand. Luckily, computers were designed to free us from these sorts of tedious tasks. The most basic tool for automating calculations is the spreadsheet program.

There are lots of spreadsheet programs including Microsoft's Excel and Apple's Numbers. Any spreadsheet program will work; they are all very similar. The instructions and screenshots here will be from Google Sheets – a free spreadsheet program you use through your web browser.

## 5.1 Your First Spreadsheet

In whatever spreadsheet program you are using, create a new spreadsheet document.

A spreadsheet is essentially a grid of cells. In each cell you can put data (like numbers or text) and formulas.



Let's put some labels in the column:

- Select the first cell (A1) and type "A number".
- Select the cell below it (A2) and type "Another number".
- Select the cell below that one (A3) and type "Their product".

- In the next column, type the number 5 in B1 and 7 in B2.

It should look like this:

	B3	
	<i>fx</i>	
	A	B
1	A number	5
2	Another number	7
3	Their product	
4		

Now put a formula in cell B3. Select B3, and type “= B1 \* B2”. The spreadsheet knows this is a formula because it starts with ‘=’. It will look like this as you type:

		<i>fx</i>	= B1 * B2
	A	B	
1	A number	5	
2	Another number	7	
3	Their product	= B1 * B2	
4			

When you press Return or Tab, the spreadsheet will remember the formula, but display its value:

		<i>fx</i>	
	A	B	
1	A number	5	
2	Another number	7	
3	Their product	35	
4			
5			

If you change the values of cell B1 or B2, the cell B3 will automatically be recalculated. Try it.

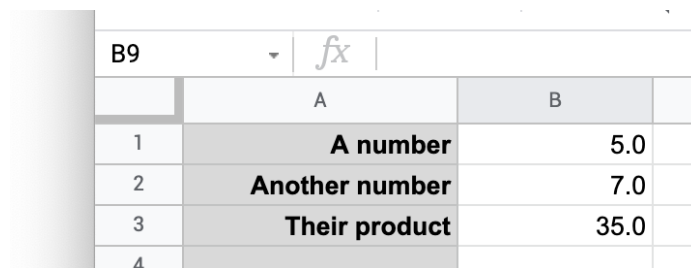


## 5.2 Formatting

Every spreadsheet lets you change the formatting of your columns and cells. They are all a little different, so play with your spreadsheet a little now. Try to do the following:

- Set the background of the first column to light gray.
- Right-justify the text in the first column.
- Make the text in the first column bold.
- Make the numbers in the second column have one digit after the decimal point.

It should look something like this:



	A	B
1	<b>A number</b>	5.0
2	<b>Another number</b>	7.0
3	<b>Their product</b>	35.0
4		

That’s a spreadsheet. You have a grid of cells. Each cell can hold a value or a formula that uses values from other cells. The cells with formulas automatically update as you edit the values in the other cells.

## 5.3 Comma-Separated Values

A lot of data is exchanged in a file format called *Comma-Separated Values* or just CSV. Each CSV file holds one table of data. It is a text file, and each line of text corresponds to one row of data in the table. The data in each column is separated by a comma. The first line of a CSV is usually the names of the columns. A CSV might look like this:

```
studentID,firstName,lastName,height,weight
1,Marvin,Sumner,260,45.3
2,Lucy,Harris,242,42.2
3,James,Boyd,261,44.2
```

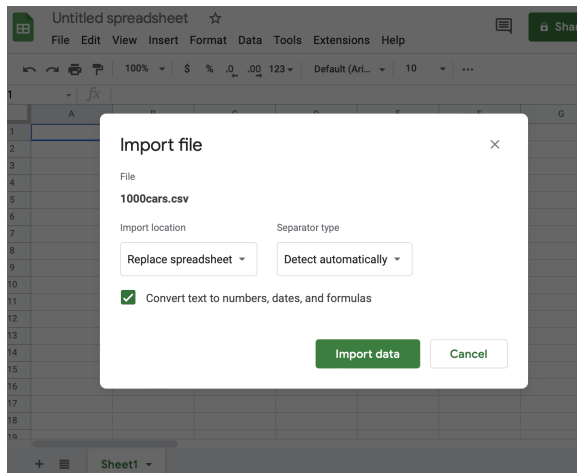
In your digital resources for this module, you should have a file called `1000cars.csv`. It is a CSV with only one column called “speed”. The first few lines look like this:

```
speed
```

33.8000  
 29.9920  
 34.8699  
 27.9936

There is a title line and 1000 data lines.

Import this CSV into your spreadsheet program. In Google Sheets, it looks like this:



You should see a long, long column of data appear. (Mine goes from cell A2 through A1001.)

	A	B	C
1	speed		
2	33.8		
3	29.992		
4	34.8699		
5	27.9936		
6	26.2875		
7	31.6701		
8	27.3347		

## 5.4 Statistics in Spreadsheets

Let's take the mean of all 1000 numbers. In cell B2, type in a label: "Mean". (Feel free to format your labels as you wish. Bolding is recommended.)

In cell C2, enter the formula `"=AVERAGE(A2:A1001)"`. When you press return, the cell will show the mean: 31.70441, if done correctly.

C8		$\sum x$	
	A	B	C
1	speed		
2	33.8	Mean	31.7044106
3	29.992		
4	34.8699		
5	27.9936		

Notice that by specifying that the function AVERAGE was to be performed on a range of cells: cells A2 through A1001.

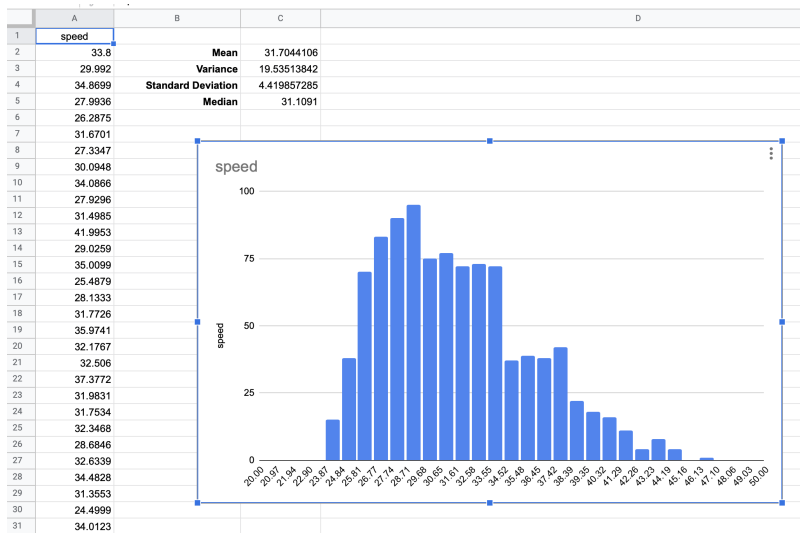
Do the calculations for variance, standard deviation, and median.

- The function for variance is VAR.
- The function for standard deviation is STDEV.
- The function for median is MEDIAN.

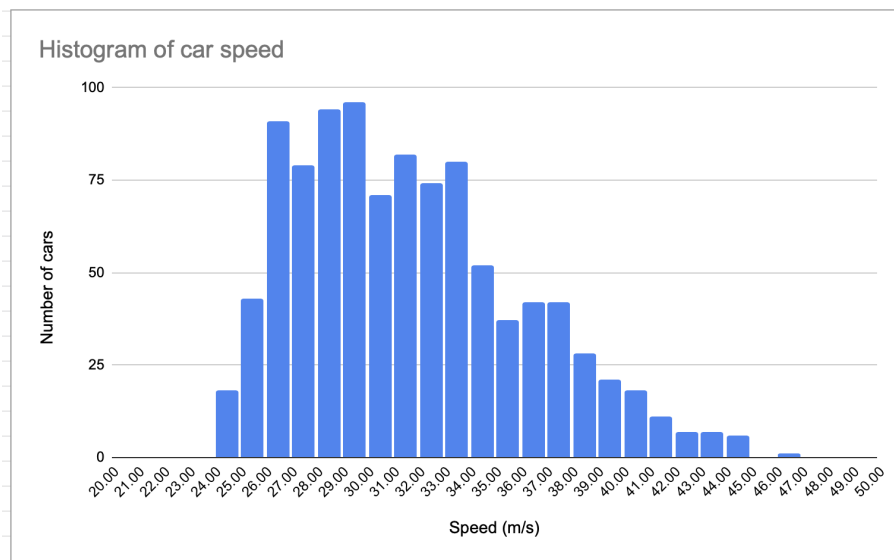
	A	B	C
1	speed		
2	33.8	Mean	31.7044106
3	29.992	Variance	19.53513842
4	34.8699	Standard Deviation	4.419857285
5	27.9936	Median	31.1091
6	26.2875		
7	31.6701		

## 5.5 Histogram

Most spreadsheets have the ability to create a histogram. In Google Sheets, you select the entire range A2:A1001 by selecting the first cell and then shift-clicking the last. Then you choose Insert→Chart. In the inspector, change the type of the chart to a histogram. This will get you a basic histogram.



Play with the formatting to see how unique you can make data. Here is an example:



**Exercise 6**      **RMS***Working Space*

In your spreadsheet, calculate the quadratic mean (the root-mean-squared) of the speeds.

You will need the following three functions:

- SUMSQ returns the sum of the squares of a range of cells.
- COUNT returns the number of cells in a range that contains numbers.
- SQRT returns the square root of a number.

*Answer on Page 41*



# Answers to Exercises

### Answer to Exercise 1 (on page 14)

Kinetic energy?  $E = mv^2 = (55)(11^2) = 6,655 \frac{\text{kgm}^2}{\text{s}^2} = 6,655 \text{ joules}$ .

Frictional force?  $F = \mu N = (0.7)(55)(9.8) = 377.3 \text{ newtons}$ .

Distance?  $D = \frac{6,655}{377.3} = 17.6 \text{ seconds}$ .

### Answer to Exercise 2 (on page 15)

The empty sled is pushing directly down on the floor with a force of  $(50)(9.8) = 490 \text{ N}$ .

The force to overcome the static friction is:

$$270 = 490\mu_s$$

Thus  $\mu_s = 0.551$

The force to match kinetic friction is:

$$220 = 490\mu_k$$

Thus  $\mu_k = 0.449$

Once you are on the sled, it is pressing directly down on the floor with a force of  $(50 + 55)(9.8) = 1,029 \text{ N}$ .

The force to overcome the static friction is:

$$F = (1,029)(0.551) = 567 \text{ N}$$

Once the sled is moving, friction is counteracting some of your force. How much?

$$F_f = (1,029)(0.449) = 462 \text{ N}$$

So all of your acceleration is due to the remaining  $567 - 492 = 75 \text{ N}$ .

We know that  $F = ma$ . In this case  $F = 75 \text{ N}$  and  $m = 105 \text{ kg}$ . So

$$a = \frac{75}{105} = 0.714 \text{ meters per second per second}$$

### Answer to Exercise 3 (on page 24)

$$\mu = \frac{1}{6} (87 + 91 + 98 + 65 + 87 + 100) = 88$$

### Answer to Exercise 4 (on page 26)

The mean of your grades is 88.

The variance, then is

$$\sigma^2 = \frac{1}{6} \left( (87 - 88)^2 + (91 - 88)^2 + (98 - 88)^2 + (61 - 88)^2 + (87 - 88)^2 + (100 - 88)^2 \right) = \frac{784}{6} = 65\frac{1}{3}$$

The standard deviation is the square root of that:  $\sigma = 8.083$  points.

### Answer to Exercise 5 (on page 27)

In order the grades are 65, 87, 87, 91, 98, 100. The middle two are 87 and 91. The mean of those is 89. (Speed trick: The mean of two numbers is the number that is half-way between.)



**Answer to Exercise 6 (on page 37)**

The formula for the RMS is “=SQRT(SUMSQ(A2:A1001)/COUNT(A2:A1001))”.





---

# INDEX

histograms, [27](#)

mean, [23](#)

median, [26](#)

quadratic mean, [28](#)

RMS, [28](#)

root-mean-squared, [28](#)

samples, [23](#)

Spreadsheet

    Entering formula, [32](#)

spreadsheet, [31](#)

summation symbol, [24](#)

variance, [24](#)