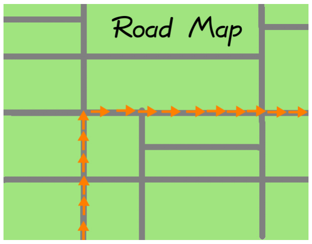
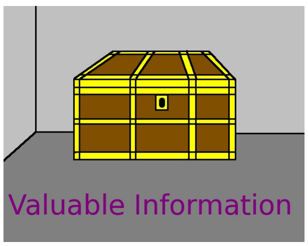
Module 2.6: Building a Linear Model



In this module we will study how one can make a model of a situation, given some data. In particular, we will focus on linear models, which are models that can be completely described by the equation of a line. This will be a very important skill in the coming modules, and we will use it often.

Some phenomena are genuinely linear, and are very faithfully modeled by linear equations. However, far more often we use a linear model not because it is flawless, but because it is both a reasonable approximation and easy to work with. Also, the skills that we study while working with linear models will be stepping stones to other types of models.



We begin with a formal definition, which might be confusing at first but which will become clear after several examples.

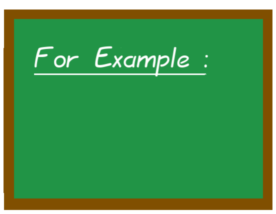
The equation of a line through the point (x_0, y_0) and with slope m is given by

$$y - y_0 = m(x - x_0)$$

Examples would include

$$y - 3 = 4(x - 5) \qquad \text{or} \qquad y + 2 = -3(x - 9) \qquad \text{or} \qquad y - 4 = 5(x + 1)$$

Furthermore, any equation in that format is said to be the equation of a line in *point-slope form*. That's because any equation written in this manner gives us an m (which is the slope) and a point (x_0, y_0) on that line.



2-6-1

We're going to warm up with some purely symbolic problems before we advance into the realm of modeling real-world situations. First, suppose someone asks you "what is the equation of a line with slope 5, and going through the point $(2, 7)$?"

The point given to us is $(2, 7)$, and so $x_0 = 2$ and $y_0 = 7$. That way we've matched the point $(2, 7)$ to the symbols (x_0, y_0) given in the previous box. Likewise, the slope was said to be 5, and so $m = 5$. Now we plug these values into the point-slope formula, and get

$$y - 7 = 5(x - 2)$$

You might be wondering how we can check our work in the previous box. Well, if we multiply out the right-hand side, we get

$$y - 7 = 5x - 10$$

and then we can add 7 to both sides and obtain

$$y = 5x - 3$$

which is in the more familiar $y = mx + b$ form or *slope-intercept form*. You can see here that the slope is 5, as requested.

While this was a useful exercise, it turns out that we can always trust the slope as provided by point-slope form, and we do not need to convert to slope-intercept form in general. The slope will be the number in front of the parentheses containing the x for any line in point-slope form.

It is much simpler to check that the equation we produced in the previous example actually makes a line going through the point $(2, 7)$. We just plug in $x = 2$ and $y = 7$, and see what comes out. From

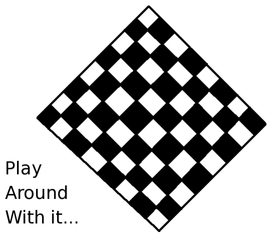
$$y - 7 = 5(x - 2)$$

we get

$$7 - 7 = 5(2 - 2)$$

and clearly what appears on either side of the equal sign is actually 0, so the equation is satisfied.

This trick—of plugging in the x and y values for x and y—will always be a useful check for any method of writing the equation of a line.

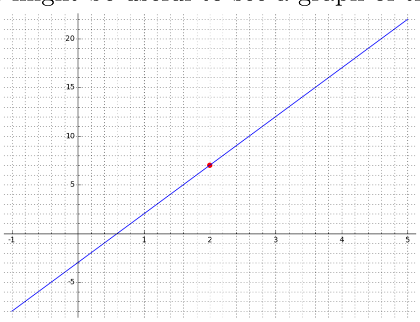


2-6-2

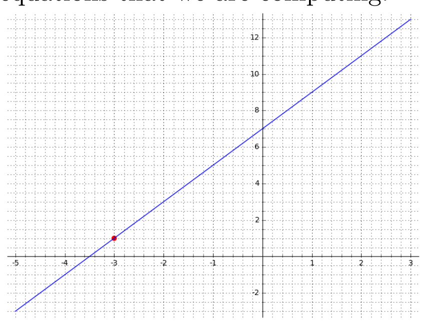
What is the equation of a line...

- With slope 2 going through the point $(-3, 1)$? [Answer: $y - 1 = 2(x + 3)$.]
- With slope 1/2 going through the point $(2, -4)$? [Answer: $y + 4 = (1/2)(x - 2)$.]

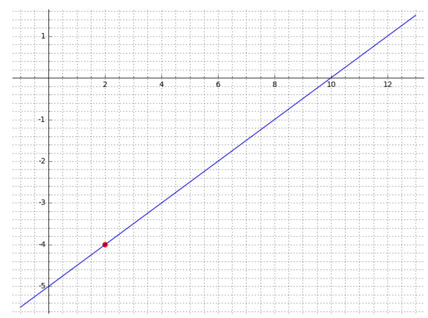
It might be useful to see a graph of the equations that we are computing.



$y - 7 = 5(x - 2)$
$-1 \leq x \leq 5$

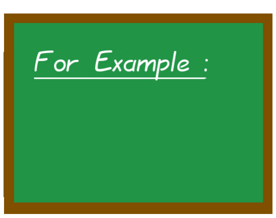


$y - 1 = 2(x + 3)$
$-5 \leq x \leq 3$



$y + 4 = \frac{1}{2}(x - 2)$
$-1 \leq x \leq 13$

As an exercise, you could figure out what you think the y-intercept should be, and then see if the graph agrees with you.



2-6-3

As we continue our warm-up before modeling real-world phenomena, let us consider the question of determining the slope of the line that connects the points $(5, 8)$ and $(3, 4)$.

You are probably familiar with the idea that slope is the change in the ys "over" the change in the xs. We can write

$$m = \frac{\Delta y}{\Delta x} = \frac{8 - 4}{5 - 3} = \frac{4}{2} = 2$$

where the triangle Δ is the Greek letter *Delta*, and means "change." In any case, the slope is 2. The symbol Δ for change is very popular—not just in math, but also in economics, science, and finance.

The utterance "the line that connects the points $(5, 8)$ and $(3, 4)$" is surely identical to the utterance "the line that connects the points $(3, 4)$ and $(5, 8)$," despite the fact that the points have changed places.

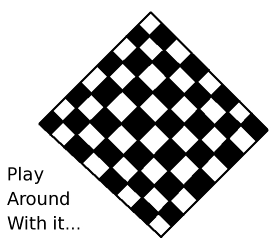
We can see this by

$$m = \frac{\Delta y}{\Delta x} = \frac{4-8}{3-5} = \frac{-4}{-2} = 2$$

which gives us the same answer.

It is important to check that the coordinates appear in the correct places in the formula for slope. In both of the previous boxes, the 4 appeared over the 3, because one of the points was $(3, 4)$. Likewise, in both of the previous boxes, the 8 appeared over the 5, because one of the points was $(5, 8)$.

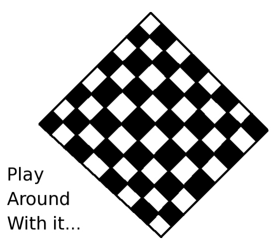
Also, it is important to remember that the ys always go in the numerator (on top), and the xs in the denominator (on the bottom). There are some scandalous memory-aids to help one recall this, involving chromosomes, but they cannot be printed here. A quick internet search should produce them, if you are curious.



2-6-4

Tell me now,

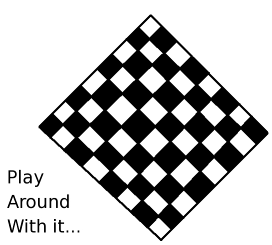
- What is the slope of the line connecting the points $(3, 8)$ and $(4, 6)$.
  [Answer: -2.]
- What is the slope of the line connecting the points $(-3, -5)$ and $(-4, -2)$?
  [Answer: -3.]
- What is the slope of the line connecting the points $(12, 14)$ and $(5, 6)$?
  [Answer: $1.14285\cdots$.]



2-6-5

Consider the points $P = (2, 10)$ and $Q = (3, 5)$.

- What is the slope of the line connecting them?
  [Answer: -5.]
- What is the equation of the line (in point-slope form) through P with slope -5?
  [Answer: $y - 10 = -5(x - 2)$.]
- What is that equation in slope-intercept form (or $y = mx + b$ form)?
  [Answer: $y = -5x + 20$.]



2-6-6

Consider again the points $P = (2, 10)$ and $Q = (3, 5)$ from the previous box.

- What is the equation of the line (in point-slope form) through Q with slope -5?
  [Answer: $y - 5 = -5(x - 3)$.]
- What is that equation in slope-intercept form (or $y = mx + b$ form, also called slope-intercept form)?
  [Answer: $y = -5x + 20$.]

The previous box gave us three equations for the line connecting P and Q. We can check them by plugging in $x = 2$ and $y = 10$, to verify that P is on that line, and by plugging in $x = 3$ and $y = 5$, to verify that Q is on that line.

Let's check the first equation, namely

$$y - 10 = -5(x - 2)$$

with

$$10 - 10 = -5(2 - 2)$$

which is true (both are 0). Also,

$$5 - 10 = -5(3 - 2)$$

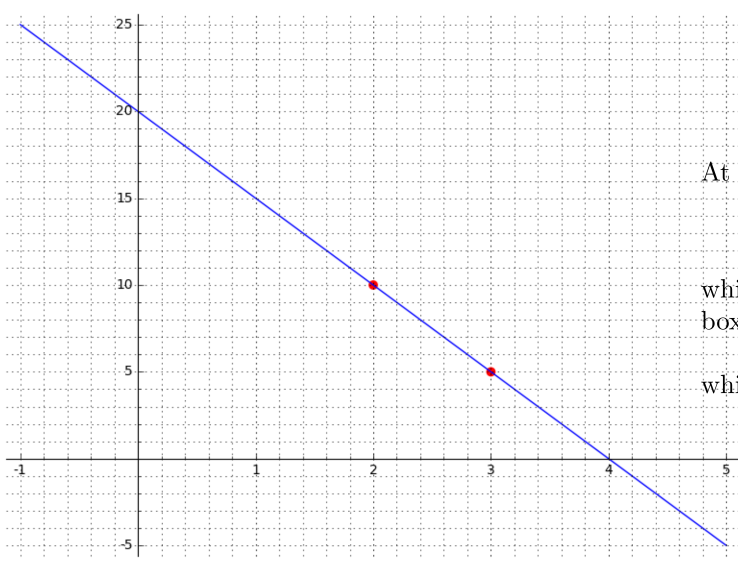
which is true (both are -5).

One could similarly check the other two equations. Why don't you please take a moment to do that now, before continuing to read on.

Sometimes students ask me if it is necessary to check both points.

Yes! You really do have to check both! Consider the potential (yet wrong) answer of $y = 5x$ for the previous box. For the point P, we have $10 = 5(2)$, and that works out fine. Yet when we check point Q, we have $5 \neq 5(3)$, and so our error is detected.

A student in a rush who checks only P and does not check Q would not detect the error, and so would be unable to fix it.

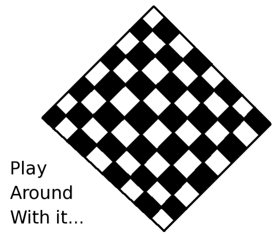


At the left, you can see the plot of

$$y = -5x + 20$$

which is the equation that we derived over the last few boxes. As you can see, I have plotted it $-1 \leq x \leq 5$.

The two red dots represent the points $(2, 10)$ and $(3, 5)$, which were P and Q.



2-6-7

Let's try this one last time. Consider the point $A = (2011, 9357)$ and the point $B = (2010, 9339)$.

- What is the slope between these two points? [Answer: 18.]
- What is the equation of a line (in point-slope form) going through the point A and having slope 18? [Answer: $y - 9357 = 18(x - 2011)$.]
- What is the equation of a line (in point-slope form) going through the point B and having slope 18? [Answer: $y - 9339 = 18(x - 2010)$.]

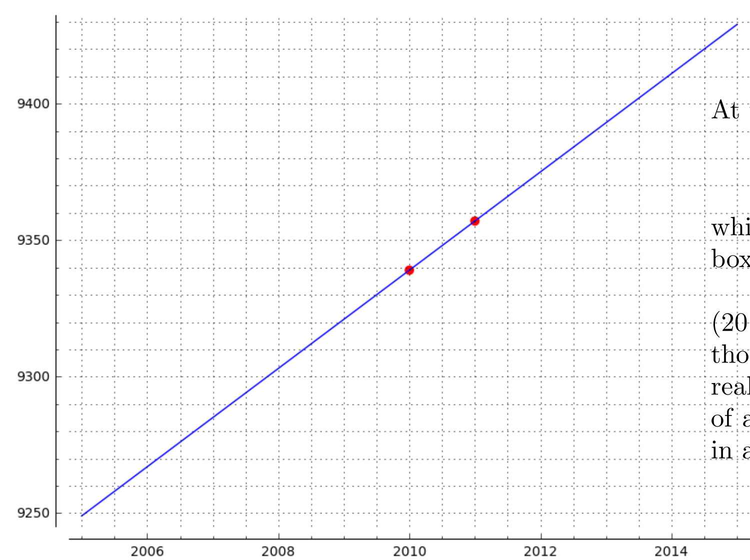
Suppose the problem in the previous box was worded "Consider the point $A = (2011, 9357)$ and the point $B = (2010, 9339)$. Find the equation of the line (in point-slope form) connecting these two points."

Perhaps you and a friend do the homework separately, and correctly compute the slope. It could turn out that you write the equation (in point-slope form) of the line through A with slope 18, whereas your friend on the other hand uses B. You would get the equation of the middle bullet in the previous box, and your friend would get the equation of the bottom bullet of the previous box. Each of you knows that you are right, by plugging in *both* Point A and Point B, but you are curious why you have two different answers. The best way to check is to convert to slope-intercept form, also known as $y = mx + b$ form. The calculation is given in the next box.

The calculations for the previous box would be

$$\begin{aligned} y - 9357 &= 18(x - 2011) \\ y - 9357 &= 18x - 36,198 \\ y &= 18x - 36,198 + 9357 \\ y &= 18x - 26,841 \end{aligned} \qquad \begin{aligned} y - 9339 &= 18(x - 2010) \\ y - 9339 &= 18x - 36,180 \\ y &= 18x - 36,180 + 9339 \\ y &= 18x - 26,841 \end{aligned}$$

As you can see, the slope-intercept equations are the same.

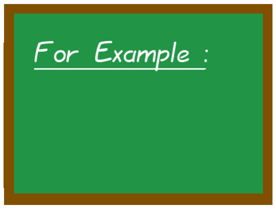


At the left, you can see the plot of

$$y = 18x - 26,841$$

which is the equation that we derived over the last few boxes. As you can see, I have plotted it $2005 \le x \le 2015$.

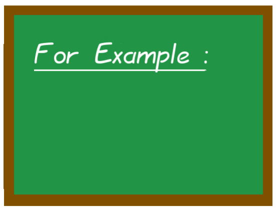
The two red dots represent the points $(2011, 9357)$ and $(2010, 9339)$, which were used to build the model. Perhaps those points look rather arbitrary or random to you. In reality, they have a very specific meaning, in the context of a real-world problem, and that will be revealed to you in a few pages.



2-6-8

We can also write linear functions instead of linear equations. Looking back at the last few checkerboards:

- Question: We were asked to find a line connecting the points $P = (2, 10)$ and $Q = (3, 5)$.
- Answer: $y = -5x + 20$ becomes $f(x) = -5x + 20$.
- Question: We were asked to find a line connecting the points $A = (2011, 9357)$ and the point $B = (2010, 9339)$.
- Answer: $y = 18x - 26,841$ becomes $f(x) = 18x - 26,841$.



2-6-9

By the way, functional notation is not restricted to $y = mx + b$ form. The previous checkerboard gave us two equations that were in point-slope form. They were

$$y - 9357 = 18(x - 2011) \qquad \text{as well as} \qquad y - 9339 = 18(x - 2010)$$

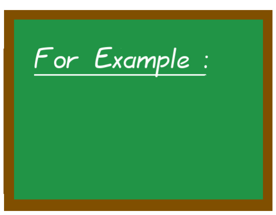
and for sure, these two can be rewritten as functions.

The only catch is that we have to move the 9357 and 9339, which are the y-intercepts, to the other side of the equal sign. Accordingly

$$y - 9357 = 18(x - 2011) \qquad \text{becomes} \qquad f(x) = 18(x - 2011) + 9357$$

and for the other one

$$y - 9339 = 18(x - 2010) \qquad \text{becomes} \qquad f(x) = 18(x - 2010) + 9339$$



2-6-10

Suppose the chancellor of your university comes to me and asks me to predict enrollment in 2012, given that the enrollment in 2010 was 9339 and the enrollment in 2011 was 9357. There are two approaches here.

First, the mathematical approach would be to build a model. We have two data points, and at this point in our exploration of this material, we have not discussed any models other than the linear model—so we must make a linear model. The year could go on the x-axis, and the enrollment on the y-axis. The model itself would be a line connecting the point $(2010, 9339)$ and the point $(2011, 9357)$. Obviously, this is exactly the problem we have been considering for the last few boxes.

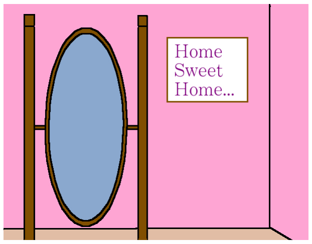
You can write either

$$y - 9357 = 18(x - 2011) \qquad \text{or} \qquad y - 9339 = 18(x - 2010)$$

and plug in 2012 for x to get an enrollment of 9375 students.

Another approach would be to say that the enrollment increased by 18 students between 2010 and 2011, and so it would not be unreasonable to assume that it would increase by 18 again between 2011 and 2012. Therefore $9357 + 18 = 9375$ is a good prediction.

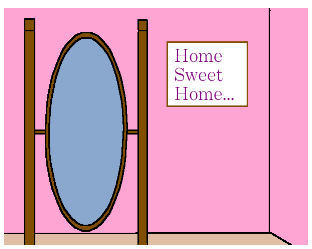
That reasoning was possible because this example is very simple—being our first modeling problem of this module. In later problems, such reasoning would be much harder, and generating a model will not only require much less work, but will also tell us additional information.



A Pause for Reflection...
There's actually something deep in the fact that we have three different equations for the line that we've been analyzing for the last several boxes. If you look at these three equations

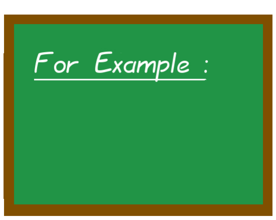
- $y - 9357 = 18(x - 2011)$
- $y - 9339 = 18(x - 2010)$
- $y = 18x - 26,841$

you might notice that important numbers from the problem (like 9357, 2011, 9339, and 2010) are found in the first two, but are absent from the third one. Admittedly, the slope, or number of new students per year, is found in all three. It is worth mentioning that the slope will always remain the same in any expression of a particular line.



A Pause for Reflection...
Even though from a mathematical point of view the three models are exactly the same model, clearly this isn't true from a real-world point of view. The first two equations are providing us easy access to certain information (like 9357, 2011, 9339, and 2010), which we cannot obtain without effort from the third equation.

The way I like to look at this, is that forms of equations are like outfits. A person can look quite different in various outfits, even though it is the same person wearing the outfits. A mathematically-oriented professional must be careful to present an equation in one form or another depending on the situation, just as one would wear different outfits to the classroom, the supermarket, the beach, or while skiing.



\# 2-6-11

We now have all the tools, but let us examine how to go smoothly from being given data, to writing down a model. Suppose a friend of yours sells ice cream from a cart in Central Park. When the price is \$ 2 per cone, he sells 200 cones; when the price is \$ 3 per cone, he sells 120 cones. Can you make a linear model that relates the number of cones sold (an economist would call this the demand) to the price?

First, we must recognize that we are asked to compute the equation of a line going through the points $C = (2, 200)$ and $D = (3, 120)$. Second, we find the slope of the line connecting C to D. We have

$$m = \frac{\Delta y}{\Delta x} = \frac{200 - 120}{2 - 3} = \frac{80}{-1} = -80$$

and then we can write either

$$y - 200 = -80(x - 2) \qquad \text{or} \qquad y - 120 = -80(x - 3)$$

We should definitely check our work before continuing on. Let's say that we use

$$y - 200 = -80(x - 2)$$

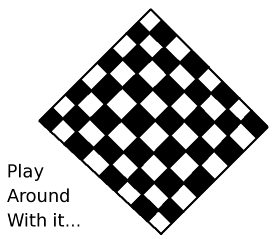
and plug in first $(2, 200)$ to obtain

$$200 - 200 = -80(2 - 2)$$

where both are 0, and so we have a match. Likewise the point $(3, 120)$ gives us

$$120 - 200 = -80(3 - 2)$$

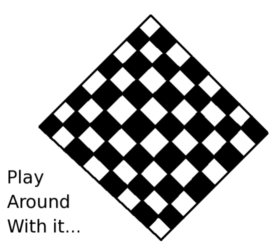
where both sides are -80, and so we have a match.



2-6-12

Now that we have this model, we can answer a few questions:

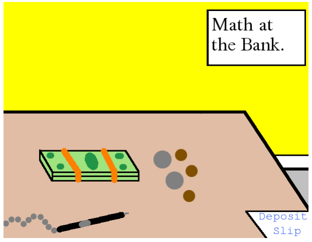
- Can you write the model $y - 200 = -80(x - 2)$ in $y = mx + b$ form?
[Answer: $y = -80x + 360$.]
- If he sells the ice-cream at \$ 1.25, how many cones does he expect to sell?
[Answer: 260 cones.]
- If he sells the ice-cream at \$ 1.75, how many cones does he expect to sell?
[Answer: 220 cones.]
- If he sells the ice-cream at \$ 2.25, how many cones does he expect to sell?
[Answer: 180 cones.]
- If he sells the ice-cream at \$ 2.75, how many cones does he expect to sell?
[Answer: 140 cones.]
- If he sells the ice-cream at \$ 3.25, how many cones does he expect to sell?
[Answer: 100 cones.]
- If he sells the ice-cream at \$ 3.75, how many cones does he expect to sell?
[Answer: 60 cones.]



2-6-13

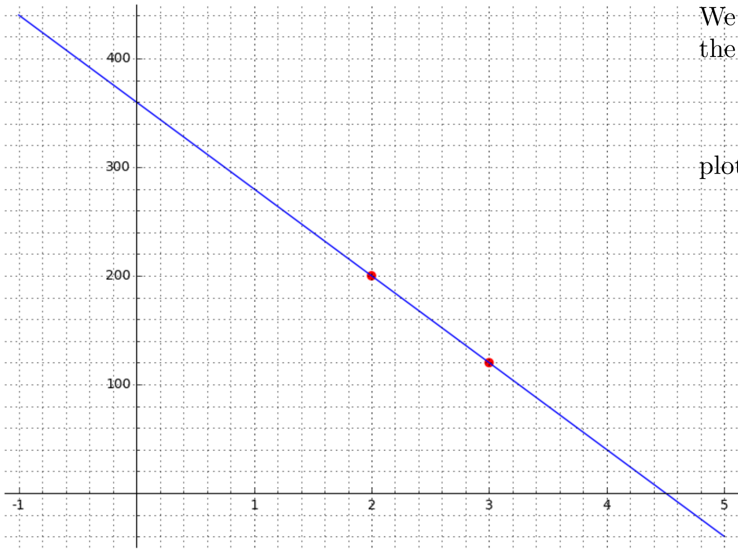
Based on the sales and prices of the previous box, compute

- What will the ice-cream salesman's total revenue be for the \$ 1.25 price level?
[Answer: \$ 325.00 per day.]
- What will the ice-cream salesman's total revenue be for the \$ 1.75 price level?
[Answer: \$ 385.00 per day.]
- What will the ice-cream salesman's total revenue be for the \$ 2.25 price level?
[Answer: \$ 405.00 per day.]
- What will the ice-cream salesman's total revenue be for the \$ 2.75 price level?
[Answer: \$ 385.00 per day.]
- What will the ice-cream salesman's total revenue be for the \$ 3.25 price level?
[Answer: \$ 325.00 per day.]
- What will the ice-cream salesman's total revenue be for the \$ 3.75 price level?
[Answer: \$ 225.00 per day.]



Of course, every businessman knows that success requires setting the price "just right," because if you set the price too high, you'll have almost no sales; if you set the price too low, then you will not get enough revenue from each sale to cover your expenses. Do you see how, in the previous box, if the price were set to \$ 2.25 per cone, our friend makes \$ 405 per day? But if he overcharges at \$ 3.25, or undercharges at \$ 1.25, he makes considerably less—namely \$ 325 per day!

We will examine this point in tremendously more detail in the next few modules, as we study "Break-Even Analysis," starting on Page 360 in the Module "Linear Break-Even Analysis."

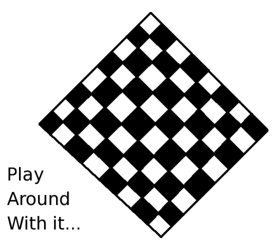


We can see here a plot of the price-demand equation for the ice cream.

$$y = -80(x-2) + 200$$

plotted from $-1 \le x \le 5$. Several facts jump out at us.

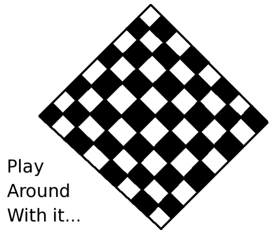
- We can see that at $x = 2$, the graph shows a sale of 200 cones, as desired.
- We can see that at $x = 3$, the graph shows a sale of 120 cones, as desired.
- We can see that at $x = 0$, the graph shows we would give away 360 free cones.
- We can see that at $x = 4.5$, the graph shows that we would sell 0 cones.



2-6-14

Suppose someone is in the business of snow removal. They have 70 customers in 2009 and 110 customers in 2011. Because the rate of increase in the number of customers is slow over time, then a linear model can be a good model.

- What would the slope of the line connecting $(2009, 70)$ and $(2011, 110)$ be? [Answer: 20.]
- What is an equation of a line with that slope going through $(2009, 70)$, in point-slope form? [Answer: $y - 70 = 20(x - 2009)$.]
- What is an equation of a line with that slope going through $(2011, 110)$, in point-slope form? [Answer: $y - 110 = 20(x - 2011)$.]
- What would these equations be in slope-intercept or $y = mx + b$ form? [Answer: $y = 20x - 40, 110$.]
- What estimate does this model give for the number of customers in 2010? [Answer: 90.]
- What estimate does this model give for the number of customers in 2012? [Answer: 130.]
- What estimate does this model give for the number of customers in 2013? [Answer: 150.]



2-6-15

Consider a successful restaurant that has recently decided to offer wine by the glass. Several test-runs are made to determine what price to charge. At \$ 9 per glass, the restaurant sells 120 glasses on the average Friday or Saturday night. However, at \$ 6 per glass, the restaurant sells 180 glasses on the average Friday or Saturday night.

- What linear model would describe how the cost of the wine affects the number of glasses bought? Let x be the price of the wine, and y the number of glasses predicted at that price, and answer in $y = mx + b$ form. [Answer: $y = -20x + 300$.]
- What would the above answer be, as a function? [Answer: $f(x) = -20x + 300$.]

There are three ways to easily check that problem's answer. We're going to plug in both given data points—namely \$ 9 per glass and \$ 6 per glass. We can also plug in their midpoint or average, which mathematicians sometimes call the arithmetic mean. The midpoint is

$$\frac{9+6}{2} = \frac{15}{2} = 7.5$$

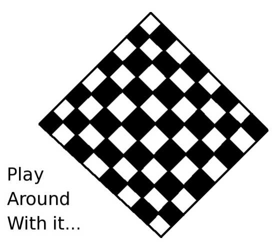
Using the two data points and the midpoint we can check

$$\begin{aligned} y &= -20(9) + 300 &= -180 + 300 &= 120 \\ y &= -20(7.5) + 300 &= -150 + 300 &= 150 \\ y &= -20(6) + 300 &= -120 + 300 &= 180 \end{aligned}$$

Now we can see that for \$ 9 per glass, and \$ 6 per glass, we get the required answers. For the midpoint, we expect

$$\frac{120 + 180}{2} = \frac{300}{2} = 150$$

and that too is a match. It will always be the case (for a linear model) that the midpoint of two inputs will produce the midpoint of the corresponding outputs.

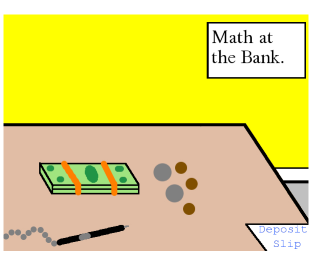


2-6-16

According to the model in the previous two boxes,

- At what price would exactly zero glasses be sold? [Answer: \$ 15 per glass.]
- How many glasses would be given out if the wine were free?

  [Answer: 300 glasses per night.]



It is nice that our model for the number of glasses sold has a specific answer for a price of \$ 0. After all, if the business owner were to make the wine available for free as a promotional for a week, there would still be some people who do not drink, or who prefer beer, as well as some who have to drive home and so cannot drink as much as they'd otherwise prefer. We will explore this point much further on, when we discuss supply and demand more formally (see Page 435 in the Module "Supply and Demand: Part One."), but for now, note that this is called a *saturation point*—the number of items that you sell when your product is free.

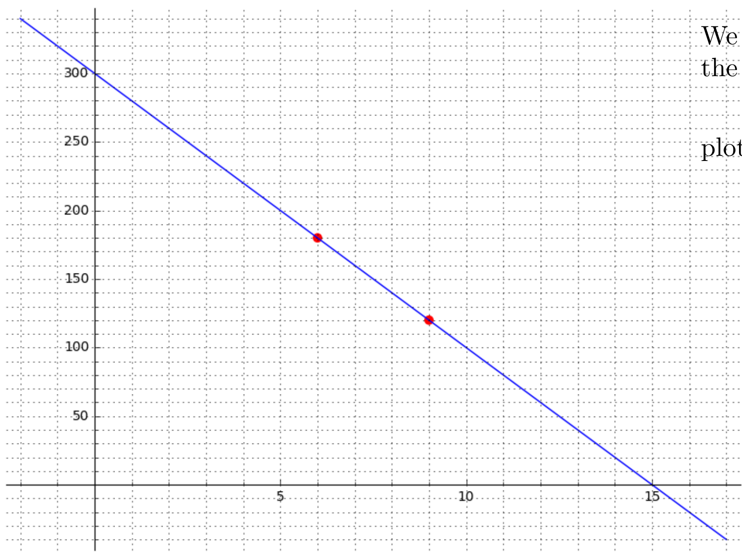
It is also good that our model for the number of glasses of wine sold will predict that at some price, 0 glasses will be sold. This is realistic after all! Would we not expect that 0 glasses would be sold, in a typical restaurant of intermediate quality, if the price per glass were \$ 1000?!

What gets interesting is if you ask how many glasses will be sold if the price is \$ 18 per glass. The model predicts:

$$y = 300 - 20(18) = 300 - 360 = -60$$

and it is difficult to imagine what "negative sixty" glasses of wine would imply.

Our model makes sense for $0 \leq x \leq 15$. We exclude $x < 0$ because a negative price makes very little sense. For $x > 15$, our model predicts that we will sell negative numbers of glasses of wine, which is weird. This interval (in which "the number sold" makes sense) is called the *domain* of the function. We will explore that topic in more detail on Page 1351.

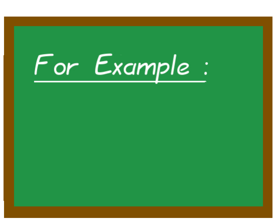


We can see here a plot of the price-demand equation for the ice cream.

$$y = -20x + 300$$

plotted from $-2 \leq x \leq 17$. Several facts jump out at us.

- We can see that at $x = 9$, the graph shows a sale of 120 glasses, as desired.
- We can see that at $x = 6$, the graph shows a sale of 180 glasses, as desired.
- We can see that at $x = 0$, the graph shows we would give away 300 free glasses.
- We can see that at $x = 15$, the graph shows that we would sell 0 glasses.



2-6-17

There is a website that computes the target heart rate for a person who is exercising, based upon their age. Fred is a physical trainer, and he would like to be able to predict the target heart rate for his clients without having to consult the website repeatedly. Therefore, he asks the website what the target heart rates are for a 24-year old, a 34-year old, and a 44-year old. The target heart rate is a zone, with an upper and lower bound. He learns that the upper bounds for these three ages are 167 beats per minute, 158 beats per minute, and 150 beats per minute, respectively. (We'll examine the lower bounds momentarily.) He is going to use the two outer data points to construct the model, and check this with the middle data point. Slope-intercept form will be of the most use for him, as it turns out.

We will solve the problem in the next box.

Continuing with the previous box, first, he must calculate the slope:

$$\frac{167-150}{24-44} = \frac{17}{-20} = -0.85$$

and then using point-slope form, with the point $(44, 150)$, we obtain

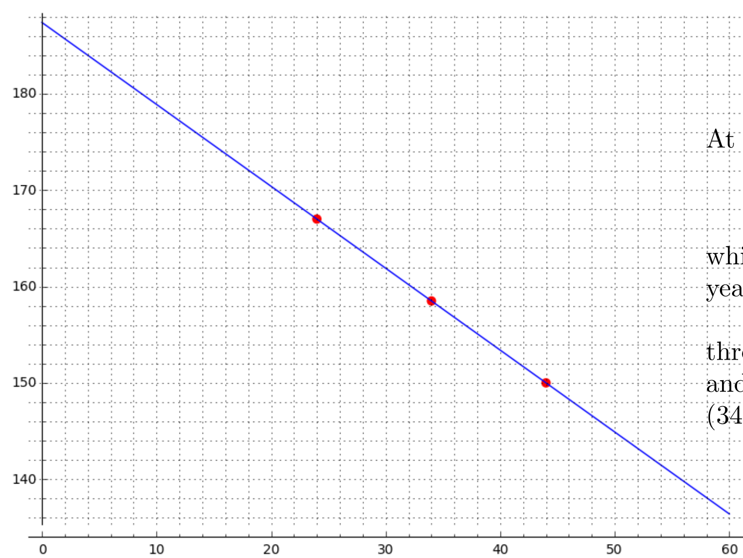
$$\begin{aligned} y - 150 &= -0.85(x-44) \\ y - 150 &= -0.85x + 37.4 \\ y &= -0.85x + 187.4 \end{aligned}$$

Last but not least, as a function this would be $f(x) = -0.85x + 187.4$.

Now we can check the work in the previous box using all three data points:

$$\begin{aligned} f(24) &= -0.85(24) + 187.4 &= -20.4 + 187.4 = 167 \\ f(34) &= -0.85(34) + 187.4 &= -28.9 + 187.4 = 158.5 \\ f(44) &= -0.85(44) + 187.4 &= -37.4 + 187.4 = 150 \end{aligned}$$

and as you can see, we have an excellent fit to the given data. We'll share the real formulas with you on Page 327 of this module.

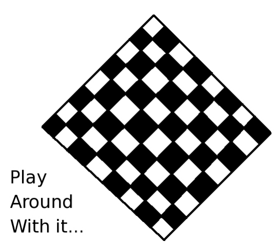


At the left you will find our plot of

$$f(x) = -0.85x + 187.4$$

which is the maximum heart rate of a person who is x years old.

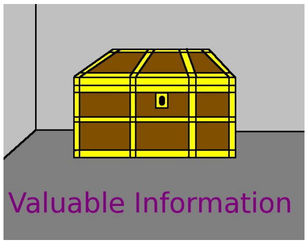
As you can see, it is plotted with $0 \leq x \leq 60$, and the three data points are marked in red. They are $(24, 167)$ and $(44, 150)$, which we used to build our model, and then $(34, 158.5)$, which helped us test our model.



2-6-18

Now it is time to examine the lower bound of the target heart rates. For a 24 year old, it is 137; for a 34 year old, it is 130; for a 44 year old, it is 123.

- Using the data of the 24-year old and the 44-year old, what is the linear model in slope-intercept $y = mx + b$ form? [Answer: $y = -0.70x + 153.8$.]
- What does this predict for the 34 year old? [Answer: 130.]



Whenever we predict anything, we should be concerned about the error of the prediction. On Page 179 of the Module "Working with Functions," we foreshadowed this topic, but now we'll introduce it properly.

The residual error, usually called "the residual," is

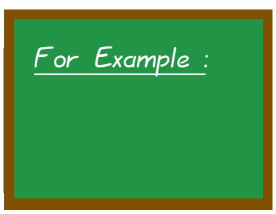
$$\text{residual} = \text{prediction} - \text{truth}$$

and the relative error can be calculated three different, but equivalent, ways:

$$\text{relative error} = \frac{\text{residual}}{\text{truth}} = \frac{\text{prediction} - \text{truth}}{\text{truth}} = \frac{\text{prediction}}{\text{truth}} - 1$$

Personally, I have always found the leftmost formula for relative error to be the easiest to understand and remember.

The residual is positive for an overestimate, and negative for an underestimate. It represents how far off the y value is. The residual would have the same units as y—for example, dollars, customers, or heartbeats. The relative error is also positive for an overestimate, and negative for an underestimate. It has no units, and can be written as a percentage by multiplying by 100. The relative error allows easy comparison among very different sorts of models (whether they involve dollars, customers, heartbeats, or anything else). Unsurprisingly, it is the preferred tool for describing the accuracy of any model.

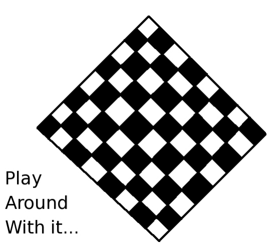


2-6-19

Suppose a magazine is transitioning from the old paper media of the 20th century to online subscribers. Let us imagine that this problem is taking place in 2001. Suppose further that they wish to forecast based on the data of 1998, 1999, and 2000, building the model based on 1998 and 2000, but checking it with 1999. They check their records and discover that there were no online subscribers in 1998, but that there were 550,000 paper-media subscribers. In 1999, there were 59,000 online subscribers, and 502,000 paper-media subscribers. In 2000, there were 122,000 online subscribers, compared with 451,000 paper-media subscribers.

Now we're going to make two linear models. The paper subscriber model, denoted $p(t)$, and the web-based subscriber model, denoted $w(t)$. To give you more practice, I'm going to permit you to find the equations of these lines yourselves before we continue.

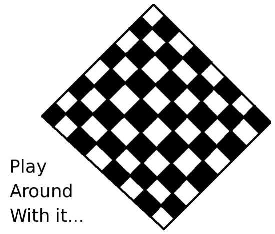
Because these problems deal with time, I'm going to make functions of t instead of functions of x. This is because t stands for time. Everything works exactly the same, where the t just plays the role of x. This is very common in mathematical economics.



2-6-20

Focusing first on the print-media subscribers, let's build the model based upon 1998 and 2000. We'll call this model $p(t)$ for "print based." Let your unit be thousands of subscribers, so that the 550,000 is handled as if it were 550.

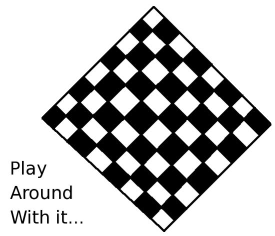
- What are the coordinates that we are using to build our model? [Answer: $(1998, 550)$ and $(2000, 451)$.]
- What is the slope of the line? [Answer: -49.5.]
- What is the function $p(t)$ in slope-intercept form? [Answer: $p(t) = -49.5t + 99,451$.]



2-6-21

Now that we have built the model $p(t)$ on the basis of the data from 1998 and 2000, let us test the model using the data from 1999.

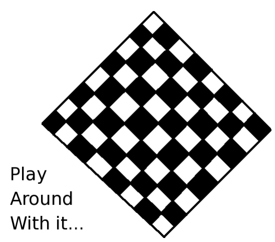
- What does this equation predict for 1999?
  [Answer: 500.5 "thousand subscribers" or 500,500 subscribers.]
- What is the residual of this approximation?
  [Answer: -1.5 "thousand subscribers" or -1500 subscribers.]
- What is the relative error of this approximation?
  [Answer: $-0.00298804\cdots$ or -0.29%. Not bad at all!]



2-6-22

Focusing now on the online subscribers, let's build the model based upon 1998 and 2000. We'll call this model $w(t)$ for "web based." Again, let your unit be thousands of subscribers, so that the 122,000 is handled as if it were 122.

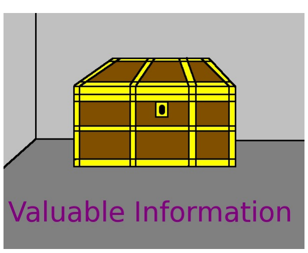
- What are the coordinates that we are using to build our model?
  [Answer: $(1998, 0)$ and $(2000, 122)$.]
- What is the slope of the line? [Answer: $+61$.]
- What is the function $w(t)$ in slope-intercept form?
  [Answer: $w(t) = +61t - 121,878$.]



2-6-23

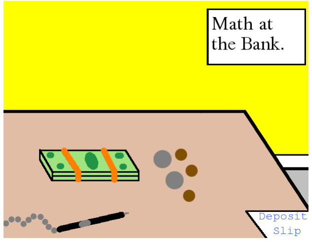
As before, because we have built the model $w(t)$ on the basis of the data from 1998 and 2000, we will test the model using the data from 1999.

- What does this equation predict for 1999?
  [Answer: 61 "thousand subscribers" or 61,000 subscribers.]
- What is the residual of this approximation?
  [Answer: $+2$ "thousand subscribers" or $+2000$ subscribers.]
- What is the relative error of this approximation?
  [Answer: $+0.0338983\cdots$ or roughly $+3.38\%$. Not bad, but not all that great!]



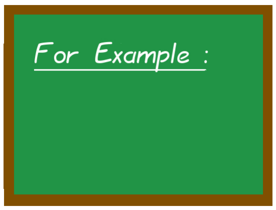
We have demonstrated a particular practice here, in both the magazine-subscription example as well as the target-heart-rate example. Namely, when given three data points, it is usually superior to build the model from the outer two data points, and check the model with the inner data point.

This is a rule of thumb and not a solid mathematical law. The reasoning behind this rule of thumb has to deal with interpolation versus extrapolation, which will be explained several times in this textbook.



The previous two boxes contain models that permit us to go further than before. We could ask, for example, when the number of online subscribers is equal to the number of print-media subscribers. We can also ask when the number of online subscribers will be double the number of print-media subscribers, or when the number of print-media subscribers is anticipated to be approximately zero.

In industry and finance, this type of forecasting can help a board of directors or CEO determine what the next few quarters are estimated to become. That, in turn, will guide their strategic decisions, and set policies for their subordinates.



2-6-24

Continuing with the problem of the last few boxes, let's start by finding out when the number of online subscribers equals the number of print-media customers. Note: we actually already solved this one, but in the context of inequalities. That was on Page 227 of the module "Review of Basic Inequalities."

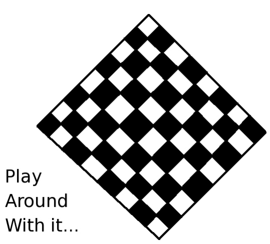
$$\begin{aligned} p(t) &= w(t) \\ -49.5t + 99{,}451 &= 61t - 121{,}878 \\ 99{,}451 &= 110.5t - 121{,}878 \\ 221{,}329 &= 110.5t \\ 221{,}329/110.5 &= t \\ 2002.97\cdots &= t \end{aligned}$$

And we learn that the number of online subscribers and print-media subscribers will be roughly equal in very late 2002, essentially the start of 2003.

We should check our work to make sure that our answer is correct. We can simply plug this point in time, 2002.97, into both models, and see if they do indeed come out equal. We obtain

$$p(t) = -49.5t + 99{,}451 = -49.5(2002.97) + 99{,}451 = -99{,}147.0\cdots + 99{,}451 = 303.985$$
$$w(t) = 61t - 121{,}878 = 61(2002.97) - 121{,}878 = 122{,}181.\cdots - 121{,}878 = 303.17$$

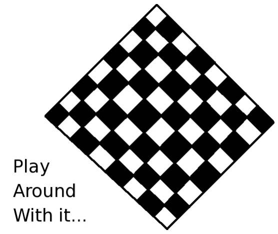
which is an almost perfect match, subjected to some rounding error. Keep in mind that these are thousands of subscribers, so roughly 303,000 of each type are expected.



2-6-25

Now we can answer some further questions about the model in the previous five boxes.

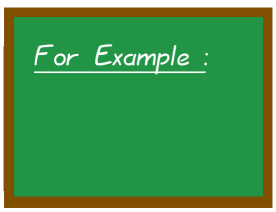
- If we want to know when the number of online subscribers will be double the number of print-media subscribers, what equation should we solve?
[Answer: $61t - 121{,}878 = 2\,(-49.5t + 99{,}451)$.]
- And what is the solution to that equation?
[Answer: $2004.87\cdots$.]
- How many subscribers of each type will there be at that time? [Answer: $419.070\cdots$ thousand or 419,070 online; and $209.935\cdots$ or 209,935 print-media subscribers.]



2-6-26

Still continuing with the previous box...

- Suppose we wish to know when the number of print-media subscribers is anticipated to be approximately zero. What equation should we solve?
  [Answer: $-49.5t + 99,451 = 0$.]
- What is the solution to that equation?
  [Answer: $t = 2009.11\cdots$.]



2-6-27

We know that the number of print-media subscribers is given by

$$p(t) = -49.5t + 99,451$$

and the number of online subscribers is given by

$$w(t) = 61t - 121,878$$

How many subscribers $b(t)$ of both types, do you predict there will be? Because $b(t) = p(t) + w(t)$, we can just do the following:

$$\begin{aligned} b(t) &= p(t) + w(t) \\ b(t) &= (-49.5t + 99,451) + (61t - 121,878) \\ &= (-49.5 + 61)t + (99,451 - 121,878) \\ &= 11.5t - 22,427. \end{aligned}$$

This is an example of adding two functions, which you probably have already read about on Page 171 of the module "Working with Functions."

The best way to check what we did in the previous box is to just plug in an arbitrary number, perhaps $t = 2009$, and see what the models tell us:

$$\begin{aligned} p(2009) &= -49.5(2009) + 99,451 = 5.500\cdots \\ w(2009) &= 61(2009) - 121,878 = 671.00\cdots \\ b(2009) &= 11.5(2009) - 22,427 = 676.500\cdots \end{aligned}$$

and sure enough

$$5.500 + 671.00 = 676.50$$

So things seem to be correct.

Depending on your handwriting, your t might look like a plus sign, or it might not. So it is possible that your variable t might be confused with a plus sign. This is not an uncommon problem, and in fact, I have it myself. This can lead to embarrassing mistakes.

The solution, however, is both simple and very effective. Just make the ts capital! You will find it is much less common that you will mistake a T for a $+$.

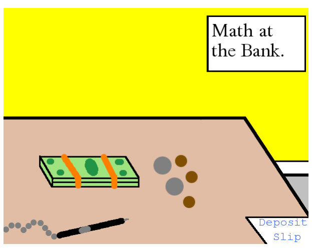
Once again, we have the issue of negative numbers doing something surprising. Because there were no online subscribers in 1998, we are not surprised to learn that if you plug in 1997 (try it) you get that $w(1997) = -61$. So our model should restrict itself to $t \geq 1998$. Likewise, you yourself calculated that $p(t) = 0$ in approximately $t = 2009.11$. So perhaps we might want to restrict our model to $t \leq 2009$.

On the one hand, if you plug $t = 2010$ into the formula for $p(t)$ you would get $p(2010) = -44$, and you'd know that something went wrong. However, on the other hand, if you plug $t = 2010$ into $b(t)$, then you will get $b(2010) = 688$, which seems very reasonable.

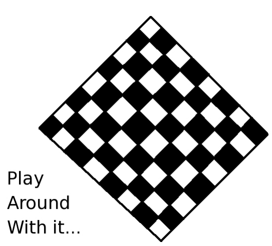
However, this $b(2010) = 688$ comes from the sum of $w(2010) = 732$ and $p(2010) = -44$ when $t = 2010$, and since a negative number of subscribers doesn't make any sense, then we have to conclude that it isn't sensible to use $t = 2010$ for $b(t)$ either.

In summary, if you build your model to work for a certain window of time, then you should not use the model outside that window of time. More formally, the sum of models should use the stricter of the two lower bounds as its lower bound, and the stricter of the two upper bounds as its upper bound.

This point is perhaps a bit too abstract in this stage of your mathematical development, but future microeconomics courses might revisit the idea.



Some of the most profound forces in manufacturing are *economies of scale*. This term refers to the fact that most things are cheaper when purchased or made in bulk. For example, a factory can afford more efficient methods of making a particular item if 10,000 of them are to be made, rather than only 100 of them. Such efficiencies are remarkably important in determining if a business plan will result in a profit or in a loss. We will explore this idea in much more detail when studying the module "Linear Break-Even Analysis," but for now let's just see an easy example.



2-6-28

Imagine yourself as a toy designer. Let's suppose that a firm is willing to manufacture a new toy at the following rates of cost: \$ 0.99 each if you manufacture 5000 of them; \$ 0.78 each if you manufacture 20,000 of them; and \$ 0.59 each if you manufacture 35,000 of them. Construct a linear model that shows the price per toy, y, for a given production order of size x, using the two outer data points, and check it with the inner data point. (Hint: Let your unit for x be "thousands of toys.")

- What two points are to be used to build the model? [Answer: $(5, 0.99)$ and $(35, 0.59)$.]
- What is the slope between those points? [Answer: -1/75 or $-0.013\overline{3}$.]
- Using the $(5, 0.99)$ data point, what is the equation of the line in point-slope form?

$$\left[\text{Answer: } y - 0.99 = (-0.013\overline{3})(x-5) \text{ or } y - 0.99 = (x-5)/(-75)\right].$$

- Using the $(35, 0.59)$ data point, what is the equation of the line in point-slope form?

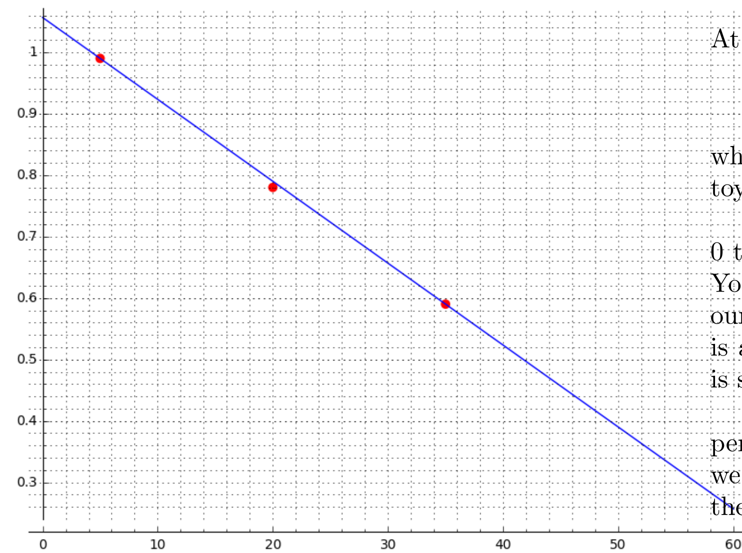
$$\left[\text{Answer: } y - 0.59 = (-0.013\overline{3})(x-35) \text{ or } y - 0.59 = (x-35)/(-75)\right].$$

- Note: normally, you'd only calculate with one point or the other point.
- What is the equation of the line in point-slope form, as a function?
[Answer: $f(x) = (-0.013\overline{3})x + 1.056\overline{6}$.]

Now let's check our work, using the two given data points and the slope-intercept form of the equation from the previous box:

$$\begin{array}{rclcl} f(5) &=& (-0.013\overline{3})(5)+1.056\overline{6} &=& -0.06\overline{6}+1.056\overline{6} &=& 0.99 \\ f(35) &=& (-0.013\overline{3})(35)+1.056\overline{6} &=& -0.46\overline{6}+1.056\overline{6} &=& 0.59 \end{array}$$

As you can see, it is a perfect fit.



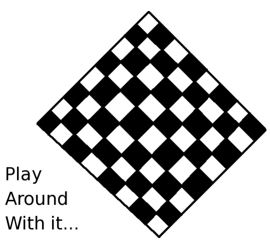
At the left we see the function

$$f(x) = (-0.013\overline{3})x + 1.056\overline{6}$$

which represents the cost per toy, when manufacturing x toys.

As you can see, I have plotted $0 \leq x \leq 60$, representing 0 toys to 60,000 toys. There are three data points in red. You can see $(5, 0.99)$ and $(35, 0.59)$, which we used to build our model, and they naturally lie solidly on the line. There is also $(20, 0.78)$, which we used to test our model. That is slightly off the line, but not by much.

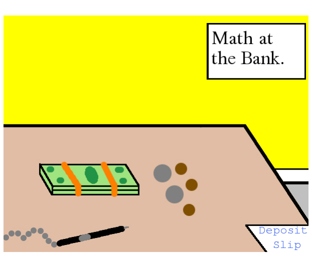
It looks as though the model is okay, definitely not perfect, but not bad. This is an imprecise assessment, so we should use some sort of mathematical tool to quantify the magnitude of the inaccuracy.



2-6-29

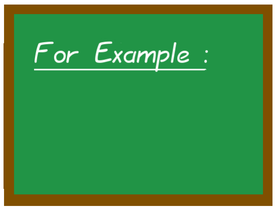
Continuing with the model of the previous two boxes,

- What does the model predict will be the case for an order of 20,000 toys? [Answer: 79 cents per toy.]
- What is the residual? [Answer: +1 cent.]
- What is the relative error? [Answer: 1.28205% or $0.012820\cdots$.]
- What does the model predict will be the case for an order of 50,000 toys? [Answer: 39 cents per toy.]



Let's suppose that a business has some important equipment. The easiest example that comes to mind is a small gym. The cardiovascular equipment might last 2, 3, or 4 years, and then it will have to be replaced, because of the fact that so many people are using it. There is a technique of accounting called *depreciation* that will accomplish two major objectives. First, it faithfully represents the reality that the equipment is worth less and less as time goes by. Second, when *depreciation* is used properly, it will ensure that a fund of money, sufficient to replace the worn-out equipment, is on hand and available for the purchasing of the new equipment. It turns out that there are many methods of deprecation. However, one of the most common is called *straight-line depreciation* and it is identical to the techniques used in this module.

The mathematical view of depreciation is a bit different from the accounting view, but they do produce the same answers when carefully applied.



2-6-30

From experience, a particular rowing machine, costing \$ 810, is known to last 36 months. Let $t = 0$ be time when the rowing machine was purchased, which was April of 2014. Let t be the time in months since that time, so that the row machine will be taken out of service when $t = 36$, which will be April of 2017. First, we want to model the changing value of the rowing machine over time with a linear function $r(t)$, from April of 2014 until April of 2017. Thus $0 \leq t \leq 36$. Second, we might have specific questions about the value of the rowing machine on certain dates.

We are really looking for a linear function $r(t)$ that has $r(0) = 810$ and $r(36) = 0$. First, we should find the slope:

$$m = \frac{810 - 0}{0 - 36} = \frac{810}{-36} = -22.5$$

and then we can find the function:

$$\begin{aligned} y - 810 &= -22.5(t - 0) \\ y - 810 &= -22.5t \\ y &= 810 - 22.5t \\ r(t) &= 810 - 22.5t \end{aligned}$$

We expect that $r(0) = 810$, because the rowing machine was worth \$ 810 when new at $t = 0$, and $r(36) = 0$, because the rowing machine will be worth \$ 0 when replaced at $t = 36$. Let's see if our function actually produces those answers.

$$\begin{aligned} r(0) &= 810 - 22.5(0) = 810 - 0 = 810 \\ r(36) &= 810 - 22.5(36) = 810 - 810 = 0 \end{aligned}$$

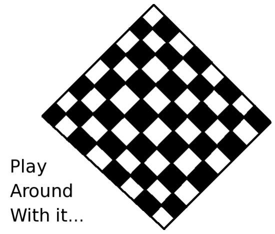
As you can see, $r(t)$ does exactly what it should.

The graph of

$$r(t) = 810 - 22.5t$$

is shown on the left. As you can see, I have graphed $0 \leq t \leq 36$, the useful life of the rowing machine. We are connecting two points, $(0, 810)$ and $(36, 0)$, with a line.

Our model of the fair market value is that it decreases, in a straight line, from \$ 810 to \$ 0 over the useful life of the rowing machine. That's why it is called straight-line depreciation.



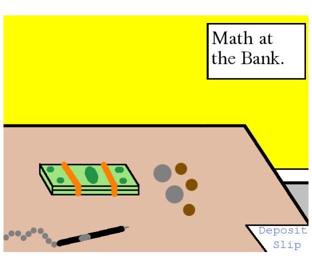
2-6-31

Let's use $r(t)$ from the previous few boxes, to answer some ancillary questions.

- We expect the mid-life value of the rowing machine to be $(810+0)/2 = 405$. Of course, that will be 18 months into the rowing machine's useful life. What is $r(18)$? [Answer: $ 405.]
- What is $r(12)$? [Answer: $ 540.]
- What is $r(24)$? [Answer: $ 270.]

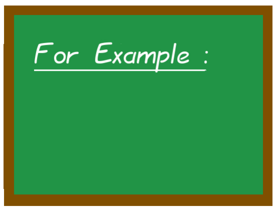
One thing that you can do with a function is make a table. You can have the input, here t, for various values. Let's try that for $f(t)$ from the last few boxes, using $t \in \{0, 1, 2, 3, \ldots, 19\}$. Then it is easy to say what $r(t)$ is for all of those values. We'll do that in the next box, and we'll toss in some additional information. We'll add $810 - r(t)$, because that represents the amount of value lost. In other words, the rowing machine is worth $r(t)$ dollars at time t, but it has lost $810 - r(t)$ dollars of value at time t.

t	$r(t)$	$810 - r(t)$		t	$r(t)$	$810 - r(t)$	
$t=0$	$ 810.00	$ 0.00	April of 2014	$t=10$	$ 585.00	$ 225.00	February of 2015
$t=1$	$ 787.50	$ 22.50	May of 2014	$t=11$	$ 562.50	$ 247.50	March of 2015
$t=2$	$ 765.00	$ 45.00	June of 2014	$t=12$	$ 540.00	$ 270.00	April of 2015
$t=3$	$ 742.50	$ 67.50	July of 2014	$t=13$	$ 517.50	$ 292.50	May of 2015
$t=4$	$ 720.00	$ 90.00	August of 2014	$t=14$	$ 495.00	$ 315.00	June of 2015
$t=5$	$ 697.50	$ 112.50	September of 2014	$t=15$	$ 472.50	$ 337.50	July of 2015
$t=6$	$ 675.00	$ 135.00	October of 2014	$t=16$	$ 450.00	$ 360.00	August of 2015
$t=7$	$ 652.50	$ 157.50	November of 2014	$t=17$	$ 427.50	$ 382.50	September of 2015
$t=8$	$ 630.00	$ 180.00	December of 2014	$t=18$	$ 405.00	$ 405.00	October of 2015
$t=9$	$ 607.50	$ 202.50	January of 2015	$t=19$	$ 382.50	$ 427.50	November of 2015



An accountant would finish their analysis with the same table that we have obtained, except without the $t=?$ column. The column that I called $r(t)$ is just called the "book value;" the column called $810 - r(t)$ is called the "accumulated depreciation." The table itself is called the "depreciation schedule."

On the other hand, the route that they would take for calculating these numbers is entirely different. They would start with $ 810, and then just repeatedly deduct $ 22.50, over and over (in MS-Excel), to complete the table. But what is $ 22.50? It is the slope of our function! (We said $m = 22.50$.) An accountant would call 22.50 the "monthly depreciation expense," instead of the slope.



2-6-32

Suppose that from experience, a treadmill is known to last five years. The new treadmill costs \$ 900. At the end of the five years, it is worth nothing for scrap value. It is purchased at the start of 2015, and will therefore be declared worn out at the start of 2020. With $t = 0$ being the start of the year 2000, what function tells me how much the treadmill is currently worth? Note, the function should only be used for years between 2015 and 2020, so $15 \leq t \leq 20$.

We are really building a linear function with the data points $(15, 900)$ and $(20, 0)$. That is because when $t = 15$, the year is 2015 and the treadmill is worth \$ 900, and when $t = 20$, the year is 2020 and the treadmill is worth \$ 0. We can find the slope:

$$m = \frac{900 - 0}{15 - 20} = \frac{900}{-5} = -180$$

and then we can easily figure out the equation of the line:

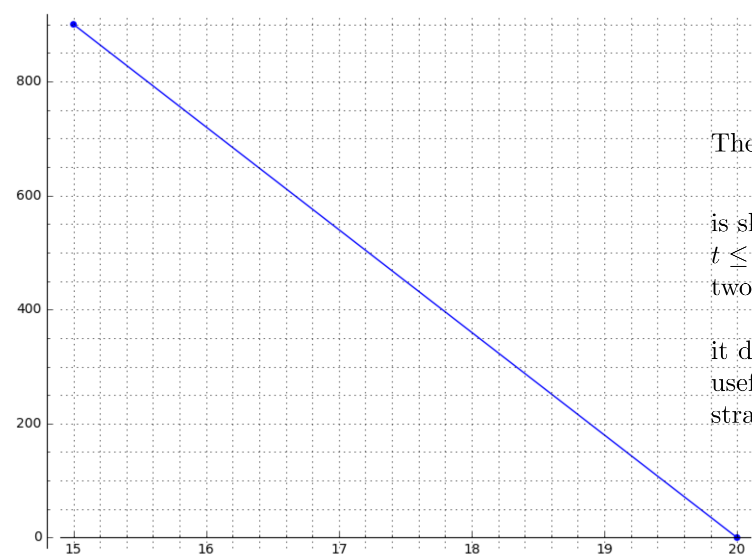
$$\begin{aligned} y - 900 &= (-180)(t - 15) \\ y - 900 &= -180t + 2700 \\ y &= -180t + 3600 \\ f(t) &= -180t + 3600 \end{aligned}$$

We will check our work in the next box.

We expect that $f(15) = 900$, because the treadmill was worth \$ 900 when new at the start of 2015, and $f(20) = 0$, because the treadmill will be worth \$ 0 when replaced at the start of 2020. Let's see if our function actually produces those answers.

$$\begin{aligned} f(15) &= -180(15) + 3600 = -2700 + 3600 = 900 \\ f(20) &= -180(20) + 3600 = -3600 + 3600 = 0 \end{aligned}$$

It is always nice when we get the answers that we expect.



The graph of

$$f(t) = -180t + 3600$$

is shown on the left. As you can see, I have graphed $15 \leq t \leq 20$, the useful life of the treadmill. We are connecting two points, $(0, 810)$ and $(36, 0)$, with a line.

As before, our model of the fair market value is that it decreases, in a straight line, from \$ 900 to \$ 0 over the useful life of the treadmill. Again, that's why it is called straight-line depreciation.

Play Around With it...

2-6-33

Let's use $f(t)$ from the previous few boxes, to answer some ancillary questions.

- We expect the mid-life value of the treadmill to be $(900+0)/2 = 450$. Of course, that will be when $t = (15+20)/2 = 17.5$, in the middle of 2017. What is $f(17.5)$? [Answer: $ 450.]
- What is $f(16)$? [Answer: $ 720.]
- What is $f(17)$? [Answer: $ 540.]
- What is $f(18)$? [Answer: $ 360.]
- What is $f(19)$? [Answer: $ 180.]

We'll construct the table (or depreciation schedule) again. I'll put the mathematical terminology at the top, and the accounting terminology at the bottom.

t	$f(t)$	$900 - f(t)$		t	$f(t)$	$900 - r(t)$	
$t = 15.0$	$ 900.00	$ 0.00	Start of 2015	$t = 18.0$	$ 360.00	$ 540.00	Start of 2018
$t = 15.5$	$ 810.00	$ 90.00	Middle of 2015	$t = 18.5$	$ 270.00	$ 630.00	Middle of 2018
$t = 16.0$	$ 720.00	$ 180.00	Start of 2016	$t = 19.0$	$ 180.00	$ 720.00	Start of 2019
$t = 16.5$	$ 630.00	$ 270.00	Middle of 2016	$t = 19.5$	$ 90.00	$ 810.00	Middle of 2019
$t = 17.0$	$ 540.00	$ 360.00	Start of 2017	$t = 20.0$	$ 0.00	$ 900.00	Start of 2020
$t = 17.5$	$ 450.00	$ 450.00	Middle of 2017				
—	book value	accumulated depreciation	year	—	book value	accumulated depreciation	year

Again, an accountant would just start with $ 900.00 and $ 0.00. Then the accountant would repeatedly subtract or add $ 90 per semiannual entry. (That matches our work, because our slope is $m = 180$ dollars per year, or $ 90 per semiannual period.) Here, the $ 90 would be called the "semiannual depreciation expense."

Play Around With it...

2-6-34

Let's suppose that an exercise bike is known to have a useful life of 36 months. Let's suppose that it costs $ 698 when new. However, it has a scrap value of $ 50, at the end of its useful life. (Perhaps a few parts can be replaced for $ 25, and then it can be sold on the used market for $ 75, or something like that.)

For simplicity, suppose that $t = 0$ when the exercise bike is bought. Clearly, the points that we are using to build our linear model shall be $(0, 650)$ and $(36, 50)$. With that in mind, compute the following:

- What is our slope (our monthly depreciation expense)? [Answer: $m = 18$ dollars per month.]
- What is the function $b(t)$, in slope-intercept form? [Answer: $b(t) = -18t + 698$.]
- Just to check our work, what is $b(0)$? [Answer: $b(0) = 698$ dollars.]
- Just to check our work, what is $b(36)$? [Answer: $b(36) = 50$ dollars.]
- Of course, at the middle of the bike's useful life ($t = 18$), we expect the bike to be worth $(698+50)/2 = 374$ dollars. So let's ask, what is $b(18)$? [Answer: $b(18) = 374$ dollars.]

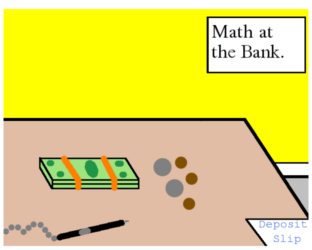
The graph of

$$b(t) = -18t + 698$$

is shown on the left. As you can see, I have graphed $0 \leq t \leq 36$, the useful life of the exercise bike.

This graph looks different from the other two. We are connecting two points, $(0, 698)$ and $(36, 50)$, with a line. Before, our function ended at \$ 0. Now, we have a scrap value (or salvage value) of \$ 50, so the function goes until \$ 50 instead. Note that the \$ 698 looks like \$ 700, because our eyes cannot see the difference.

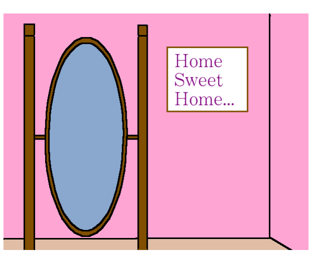
We can say that our model of the fair market value is that it decreases, in a straight line, from \$ 698 to \$ 50 over the useful life of the exercise bike.



By the way, it is very common to consider "scrap value," sometimes called "salvage value," during depreciation calculations in the actual business world.

Meanwhile, another interesting distinction is "fair market value" versus "book value." The numbers that you get from depreciation calculations are the "book value." That's what you write into the company's financial records (called books in the slang of accounting). The book value, being a result of our calculations, is our model of the fair market value.

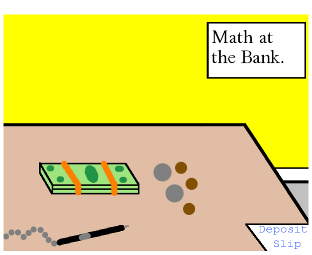
The fair market value is what the exercise bike, treadmill, or rowing-machine would actually sell for in an auction or a used equipment sale. In truth, it might not change in a straight line. For example, lots of people start exercising right after New Year's Day, or in late spring. The price would be slightly higher if you sold an exercise bike on January 2nd than on some random day late in the year. Because the fair market value is somewhat difficult to predict, we really do need two different vocabulary terms. After all, the book value is the result of a formula, and it is absolutely certain.



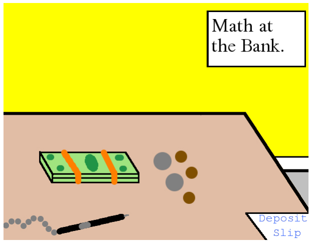
A Pause for Reflection...

Now I have shown you both the accountant's approach and the mathematician's approach to straight-line depreciation. As you can see, it is a very easy concept. Most of my students find the account's method easier, and I am not sure that I disagree. However, many pieces of factory equipment last 40 years, and it would be tedious to make a table (a depreciation schedule) with 40 rows, for an annual model, or 480 rows, for a monthly model. Making a mathematical function would be far more efficient in that case.

In any case, I hope it was useful for you to compare and contrast the two methods.

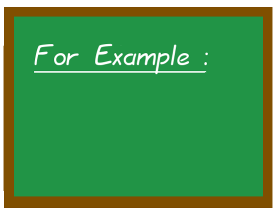


There are other methods of depreciation besides straight-line depreciation. The "declining balance" and "double-declining balance" methods are preferred for computers and other digital electronics, which rapidly lose value soon after purchase. The graph produced by such methods is a curve, and not a line. This represents the fact that the depreciation is very rapid at the start of the item's useful life, but somewhat slower in later years. We won't go into that in this textbook.



Another method is the "units of production" method, and that's great for factory equipment. For example, a particular machine might produce one million jet-engine blades during its useful life. The number of blades produced to date is used as the depreciation variable, not time. This is sometimes used for taxicabs, where the number of miles driven becomes the variable, instead of time.

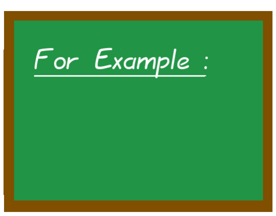
Last but not least, the US Federal government uses something called MACRS (Modified Accelerated Cost Recovery System) for tax purposes. Tax regulations are complicated. For example, if the scrap value or salvage value is less than 10% of the total cost, then you can pretend that the scrap/salvage value is zero. However, if it is 10% of the total cost or more, then you must include it in the calculation. This stuff can certainly get complicated, and I'm sure we can all understand why accountants are very well paid.



2-6-35

The slope of a line is very evident from either the point-slope form or the slope-intercept form. Moreover, when making a linear model, often the slope of the line can tell us additional information. In practice, that information is usually quite interesting. Let's consider some examples that we have seen earlier.

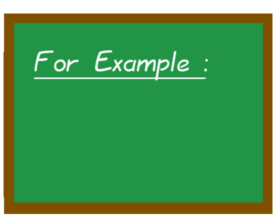
- In the student enrollment example, the slope of 18 represented an increase of 18 students per year in the university's enrollment.
- In the ice-cream salesman example, the slope of -80 represented an decrease of demand of 80 cones per dollar increase in price, or equivalently, -40 cones per 50 cent increase in price. This is most evident looking at the answers in the "checkerboard" box immediately after the example.



2-6-36

Let's consider the role of the slope in the analysis of the snow-removal business.

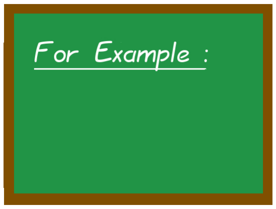
The slope of 20 represented a *net* increase in 20 customers per year. However, this does not necessarily mean that no one quits the service and you have 20 new customers—though that is possible. It could mean 25 new customers signed up while 5 customers quit using the service—this is also possible. Perhaps 29 new customers signed up while 9 customers stopped the service—that too is possible. All three possibilities and many others would have the slope exactly at 20.



2-6-37

We should also consider the role of the slope in the wine-by-the-glass example.

The slope was -20, and this indicates that each dollar increase in the price per glass will cause you to lose 20 glasses worth of wine sales per average Friday or Saturday night. Alternatively, each 50-cent increase in the price per glass will cause you to lose 10 glasses worth of wine sales per average Friday or Saturday night.

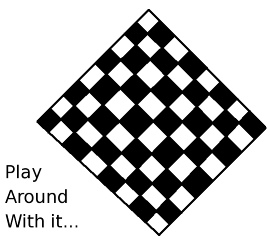


2-6-38

Let's review the role of the slope in the three depreciation models.

- For the rowing machine, we had $r(t) = 810 - 22.5t$, which indicated that the book value decreased by \$ 22.50 per month, and accumulated depreciation increased by \$ 22.50 per month. This is the monthly depreciation expense.
- For the treadmill, we had $f(t) = -180t + 3600$, which indicated that the book value decreased by \$ 180 per year, and accumulated depreciation increased by \$ 180 per year. This is the annual depreciation expense.
- For the exercise bike, we had $b(t) = -18t + 698$, which indicated that the book value decreased by \$ 18 per month, and accumulated depreciation increased by \$ 18 per month. This is the monthly depreciation expense.

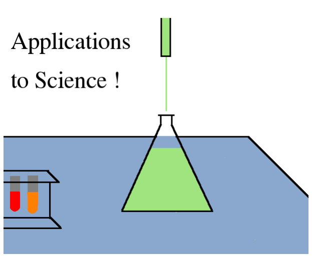
As you can see, the slope is extremely relevant in depreciation models, because it tells us the rate at which the equipment is losing value.



2-6-39

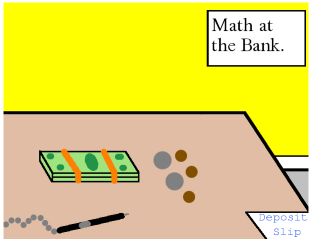
Consider the newspaper that was transitioning from being print-based to being internet-based. (This problem began on Page 307 of this module.) We had three linear models there.

- Write a sentence or phrase to describe the meaning of the slope for $w(t)$. As a bonus, can you tell us why it was positive?
- Write a sentence or phrase to describe the meaning of the slope for $p(t)$. As a bonus, can you tell us why it was negative?
- Write a sentence or phrase to describe the meaning of the slope for $b(t)$. As a bonus, can you tell us why it was positive?
- Keep in mind that because the answers are sentences or phrases, you will be very unlikely to produce exactly the same wording as I do. However, so long as you are relatively close, then that's okay.
- The answers to these questions can be found on Page 328 of this module.



The slope in our medical examples can also tell us something interesting. Reconsider the ideal heart rate problem on Page 305 of this module.

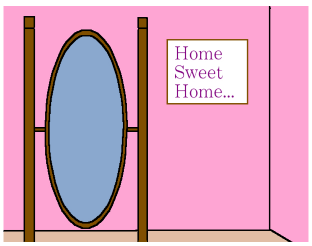
- Here we had slope -0.85 for the upper bound, and slope -0.70 for the lower bound.
- The negativity of the slope implies that as you age, you should expose your heart to rates that are somewhat lower than what you are accustomed to at a younger age. In particular, every 10 years that you age, you should change your threshold by 8.5 beats per second for the upper bound, and 7 beats per second for the lower bound.
- We will see further implications of the -0.85 and -0.70 later, on Page 327 of this module.



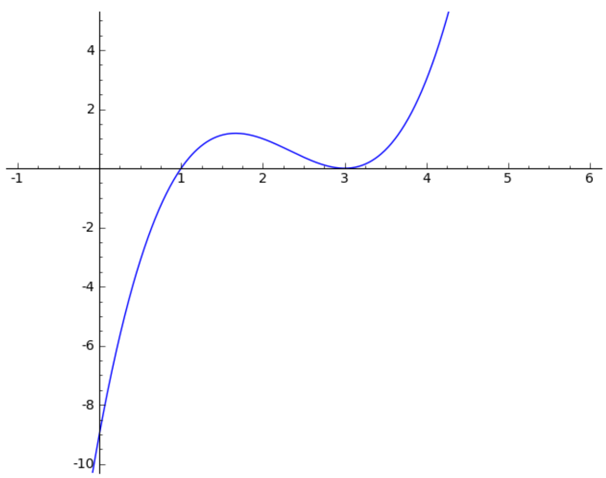
The slope of the economy-of-scale example (from Page 311 of this module) is a foreshadowing of a point that you will study in greater detail and in more realism in an advanced economics course, provided that you choose to take one.

By a slope of $-0.013\overline{3}$, what the model is telling you is that every time x goes up by 1, the y goes down by $0.013\overline{3}$. So what we should write is "every time that we increase production by one toy per batch, the price of the batch drops by $0.013\overline{3}$ cents. While that is an excellent example of mathematical precision, it is too wordy to be understood by many people. Therefore, we could also write "every time you increase production by 1000 toys, the cost per toy drops by $13.3\overline{3} = 13\frac{1}{3}$ cents each."

Basically, it is often the case that a large factory (producing many more items) will have a lower per-item cost than a smaller factory. The slope allows you to see precisely how much.



A Pause for Reflection. . .
When looking at each of those cases the slope told us something more or less valuable, as a quick review of the previous boxes will tell us. In each specific problem, the idea of slope conveyed a significant information and can be thought of as relatively important. In the overall view, taking all the problems together, the notion of slope is an indescribably useful and powerful tool.

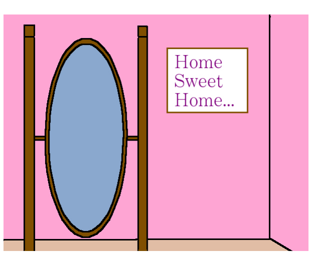


Obviously linear models have slopes, as we've explored this to a great extend already. Now look at the graph to the left of this box, which is clearly not a line. In fact, the graph to the left is the equation

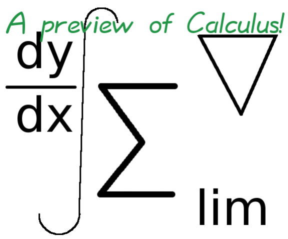
$$y = (x-1)(x-3)^2$$

which is a nice example of a cubic polynomial. Now you can see that between 0 and 1 the y-value is increasing very rapidly. This would imply a large positive slope. Between 2 and 3, however, the y-value is actually decreasing, which would imply a negative slope. Between 1.5 and 2, the y-value is not changing all that much, and that would imply that the slope is somewhat near zero, whether it be positive or negative.

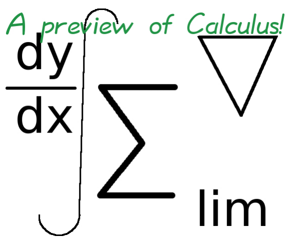
Therefore, you can see that no single number alone could possibly describe the slope of this curve at all x-values shown. Instead, we would need something more complicated than a mere number.



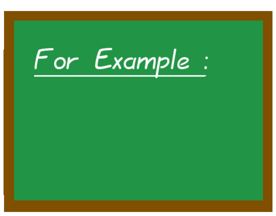
A Pause for Reflection. . .
Now seeing how useful the slope is as a concept, but also seeing how it clearly does not work for certain types of equations like cubics, perhaps you might imagine if there is an opportunity here. Wouldn't it be nice if someone were to invent an analog of slope, that works for equations more complex than lines?



The idea of the previous box is precisely what the co-inventors of calculus discovered. Isaac Newton (1642–1727) and Gottfried Wilhelm Leibniz (1646–1716) created the theory of "derivatives," which describe a rate of change, like a slope does. The difference is that the slope of a cubic or a parabola or anything other than a line is not a number, but a function. In any case, calculus permits a mathematician to apply the concept of a rate of change to things other than lines. In that sense, calculus can be said to be "the study of rates of change."



In fact, there are only four pieces to calculus. The most important is the derivative, which I've just told you about. Another piece is called "the integral," and simply starts with a derivative and returns back to you the original model producing that derivative. The two other pieces are called "limits" and "sequences."



2-6-40

Having throughly analyzed the cost of manufacturing batches of toys from the problem on Page 311 of this module, we are able now to compute the total cost of the entire batch. If the toys cost $f(x)$ dollars each to manufacture, and we are manufacturing x of toys, then we are going to pay $xf(x)$ dollars in total for the entire batch. However, it might be nice to have one function that produces a cost right away. Therefore, we should seek out a function $g(x) = xf(x)$, which would produce for us the total cost.

We recall

$$f(x) = (-0.013\overline{3})x + 1.056\overline{6}$$

was our function, but x is measured in thousands. We will find it easier to convert it to have x measured in toys. We have then

$$f(x) = (-0.000013\overline{3})x + 1.056\overline{6}$$

where we simply divided the coefficient of x by 1000.

We will continue in the next box.

Continuing with the previous box, we can confidently say that a batch of x toys would cost $g(x) = xf(x)$ dollars. We obtain the following:

$$\begin{aligned} g(x) &= xf(x) \\ &= (x)\left((-0.000013\overline{3})x + 1.056\overline{6}\right) \\ &= -0.000013\overline{3}x^2 + 1.056\overline{6}x \\ &= -(1.3\overline{3} \times 10^{-5})x^2 + 1.056\overline{6}x \end{aligned}$$

This idea will become important when we study "non-linear break-even analysis," which begins on Page 1212.

Let's check our work. One of the original data points from the question which gave birth to this model is that if we make 5000 toys, the cost will be 99 cents each. Let's use that to compare the model to common sense. If we have 5000 toys at 99 cents each, we would expect

$$5000 \times 0.99 = 4950$$

dollars in total costs. Meanwhile, the model says for $g(5000)$ that we should have the following:

$$\begin{aligned} g(x) &= -(1.3 \times 10^{-5})x^2 + (1.056\overline{6})x \\ g(5000) &= -(1.3 \times 10^{-5})(5000)^2 + (1.056\overline{6})(5000) \\ g(5000) &= -(1.3 \times 10^{-5})(25,000,000)^2 + (1.056\overline{6})(5000) \\ g(5000) &= -333.3\overline{3} + 5283.3\overline{3} \\ g(5000) &= 4950.00 \end{aligned}$$

Continuing with the previous box, we can see what happens for 35,000 toys. We expect those toys to be 59 cents each, so the total cost should be

$$(35,000)(0.59) = 20,650.00$$

Meanwhile, the model says for $g(35,000)$ that we should have the following:

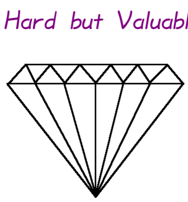
$$\begin{aligned} g(x) &= -(1.3 \times 10^{-5})x^2 + (1.056\overline{6})x \\ g(35,000) &= -(1.3 \times 10^{-5})(35,000)^2 + (1.056\overline{6})(35,000) \\ g(35,000) &= -(1.3 \times 10^{-5})(1,225,000,000)^2 + (1.056\overline{6})(35,000) \\ g(35,000) &= -16,333.\overline{3} + 36,983.\overline{3} \\ g(35,000) &= 20,650.00 \end{aligned}$$

What happens if we try to manufacture $x = 1,000,000$ toys? What would we get?

$$\begin{aligned} g(x) &= -(1.3\overline{3} \times 10^{-5})x^2 + (1.056\overline{6})x \\ g(1,000,000) &= -(1.3\overline{3} \times 10^{-5})(1,000,000)^2 + (1.056\overline{6})(1,000,000) \\ g(1,000,000) &= -(1.3\overline{3} \times 10^{-5})(10^{12}) + 1,056,666.\overline{6} \\ g(1,000,000) &= -1.3\overline{3} \times 10^{7} + 1,056,666.\overline{6} \\ g(1,000,000) &= -1.22276\overline{6} \times 10^{7} \end{aligned}$$

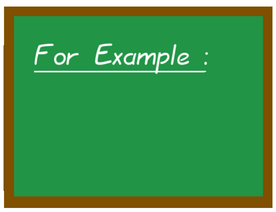
This is patently absurd. Surely you would not have negative costs. What has gone wrong? Our original model of $f(x) = (-0.000013\overline{3})x + 1.056\overline{6}$ was build on the x-values of 5, 20, and 35, representing 5000 to 35,000 toys. The value of $x = 1000$ to represent a million toys is way out of range. We are using the model to discuss something it was not built to discuss. Therefore it gives us a wrong answer. This wrongness continues when we try to use $g(x)$, because we built $g(x)$ upon $f(x)$.

This box is yet another example of *extrapolation*, something which you should avoid at all costs.



In the Fall of 2010, I had an excellent student from the People's Republic of China, Zhao Yang; he taught me how they approach problems of the "building a linear model" category in China. As it comes to pass, I think the method is a bit harder, but other students disagree. I thought it might be nice to include this approach, which we will explore in the next three boxes. Some instructors may wish to skip this material. You should check with your instructor before proceeding.

After that, I'd like to introduce two more concepts. The first is the idea of turning a function inside-out. The second is a formula, called the "point-point form of a line," which is hard to remember. However, this hard-to-remember formula can build a linear model very quickly. I am told that it is somewhat popular among biologists.

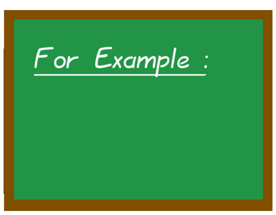


2-6-41

We return to the magazine subscription problem from Page 307 of this module. In 1998 there were 550,000 paper subscribers, and in 2000 there were 451,000 paper subscribers. The two points are thus (1998, 550) and (2000, 451). We then have

$$\begin{array}{lrcl} \text{Eqn. \#1:} & 550 & = & m(1998) + b \\ \text{Eqn. \#2:} & 451 & = & m(2000) + b \end{array}$$

which is merely a pair of linear equations in two variables, something we know quite well how to solve. We'll solve them in the next box.



2-6-42

Continuing with the previous box, we're going to solve the two equations given to get m and b. Subtracting those equations yields

$$(550 - 451) = m(1998 - 2000) + b(1 - 1)$$

or more plainly

$$99 = -2m$$

from which we learn that $m = 99/(-2) = -49.5$ and we plug m back into the first original equation. The calculation continues in the next box.

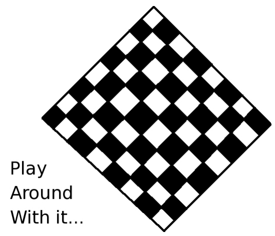
Continuing with the previous box,

$$\begin{array}{rcl} 550 & = & (-49.5)(1998) + b \\ 550 & = & -98,901 + b \\ 550 + 98,901 & = & b \\ 99,451 & = & b \end{array}$$

giving us the answer

$$y = -49.5x + 99,451$$

which is, of course, exactly the same as what we got on Page 307 of this module. Whether this is easier or harder is a matter of personal taste. I find that different students have extremely different opinions on what they find easier or harder.



2-6-43

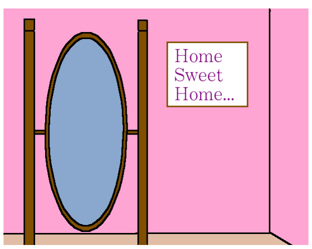
Consider what equations you might write to use this Chinese method for the student enrollment problem from Page 300 of this module.

$$\left[\text{Answer:} \quad \begin{array}{rcl} 9339 & = & m(2010) + b \\ 9357 & = & m(2011) + b \end{array}\right]$$

Now solve the equations, and see if your model matches what we obtained, namely

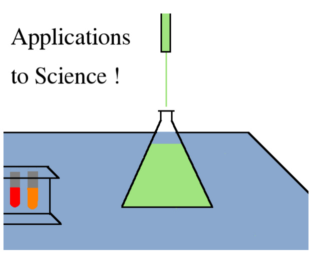
$$y = 18x - 26,841$$

In the next four boxes, we're going to discuss the reliability of the predictions that we have computed in this module.



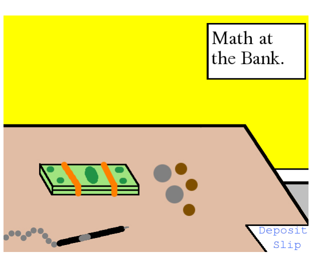
A Pause for Reflection...
The problems on student enrollment, the cost of ice-cream, wine-by-the-glass, and magazine subscriptions all have something in common: they rely on human nature. In particular, humans can be picky or finicky about what they buy. The demand of ice-cream can vary by the weather; wine-by-the-glass sales could be influenced by fads in alcoholic beverages; the magazine subscriptions could vary depending on the content of the magazine and the economy. Finally, student enrollment could be influenced by the university's ratings, changing student loan conditions, and the economy.

Since human factors come in to play, and since no one can predict if any individual human will buy a glass of wine or an ice cream cone, these models are of intermediate utility only. They certainly aren't useless—in fact these techniques are taught in economics courses worldwide. Yet, human desires—and most particularly the demand for a product—are notoriously hard to predict mathematically.



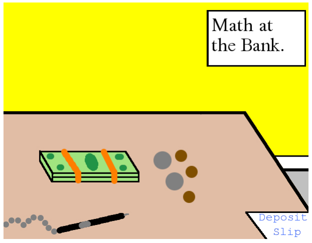
In contrast, the ideal-heart-rate model is based on medical evidence. The age and the heart rate are are measurable numbers. On the other hand, the "ideal" is set by the medical community, and while this is hopefully based on solid evidence, it could change if some startling study reveals that the model is unsound. That would be a surprise, however, and is unlikely.

These scientific models are much more reliable than models of human desire or activity.

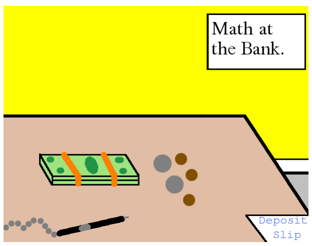


In the previous two boxes, you may have noticed that we did not characterize the snow-removal problem and the economy-of-scale problem related to the manufacturing of toys. These form a middle ground.

While the weather can be good one year, and bad another year, resulting in varying demand for snow removal, few businesses or residences would opt out of having their parking lot plowed after a major storm—particularly once the size gets above what a single person can do with a shovel by himself or herself. Thus the data can change but since businesses truly need to provide parking to customers, the demand is less subject to human taste. One does not genuinely need a glass of wine or an ice cream in the same true sense of "need."



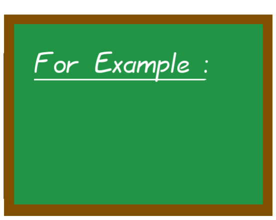
The previous box is an example of the important and fascinating economic concept of *elasticity of demand*, which sadly we cannot explore deeper in this book. Suffice it to say that *inelastic goods* are those whose purchase by customers is not optional (like food and medicine). While this might be unpleasant in times of high prices, it does result in the mathematical modeling of the demand for these goods becoming more exact. Conversely, goods whose purchase is very optional (like wine and ice cream) are *elastic goods*. Those have demand that is harder to predict by building a linear model, but the method of "Supply and Demand," also covered in this book, does well for elastic goods.



In the toy manufacturing problem, the factory's underlying costs would be negotiated in a contract, and that contract would be unable to change without mutual consent. During the window of time the contract is in force, the price structure must not change. The mathematics is entirely trustworthy during that window. At the end of that window, however, renegotiations often take place, and the role of human nature returns, severing the role of mathematics in the problem until a new contract is signed.

Well, up to this point in the module, we've been focusing on finding equations that relate x and y, and we have also found some nicer forms that can compute y instantly when x is known.

What about the reverse direction? What if you know y and want to know x rapidly? Is there a form for that? Yes! There is! You could call this "turning a linear model inside-out," and it will (much later) lead to the topic of "inverse functions."



2-6-44

On Page 304 of this module, we calculated that if the wine was priced at x dollars a glass, we would sell y glasses, where $y = 300 - 20x$. This is a useful formula when we know the price, but want to know the demand. On the other hand, how might we come up with a similar formula, for what price we should charge to produce a given demand?

What we're going to do is to solve this equation for x.

$$\begin{aligned} y &= 300 - 20x \\ y - 300 &= -20x \\ \frac{y-300}{-20} &= x \\ \frac{300-y}{20} &= x \\ 15 - y/20 &= x \end{aligned}$$

While the last formula is the most compact, any of the last three lines there would be a correct answer.

Since we know what prices will result in sales of 120 glasses, 150 glasses, and 180 glasses, then we can use these as test values:

$$\begin{aligned} x &= 15 - 120/20 &= 15 - 6 &= 9.00 \\ x &= 15 - 150/20 &= 15 - 7.5 &= 7.50 \\ x &= 15 - 180/20 &= 15 - 9 &= 6.00 \end{aligned}$$

and we see that these come out to the prices that we had found on Page 304 of this module.

Play Around With it...

2-6-45

Let's try some symbolic problems to practice this skill.

- If $y = 3x + 60$ then what is x in terms of y? [Answer: $x = (y/3) - 20$.]
- If $y = 5x + 120$ then what is x in terms of y? [Answer: $x = (y/5) - 24$.]

Play Around With it...

2-6-46

On Page 301 of this module, you learned that if the price of ice cream is x, then our vendor can expect to sell y ice-cream cones, where

$$y = -80x + 360$$

That being the case, write an equation that gives x in terms of y.
[Answer: $x = 4.5 - y/80$.]

Play Around With it...

2-6-47

On Page 311 of this module, you learned that if a toy manufacturer produces x toys (measured in thousands), then the cost of manufacturing each toy is y dollars, where

$$y = (-0.013\overline{3})x + 1.056\overline{6}$$

That being the case, write an equation that gives x in terms of y.
[Answer: $x = 79.25 - 75y$.]

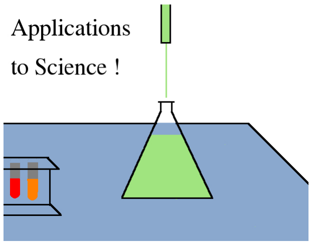
Play Around With it...

2-6-48

On Page 306 of this module, you learned that if someone is x years old, then the lower bound of their ideal heart rate (in beats per minute) is given by y where

$$y = -0.70x + 153.8$$

That being the case, write an equation that gives x in terms of y.
[Answer: $y = 219.714\cdots - (1.42857\cdots)x$.]



If you recall, when we first introduced the target-heart-rate information on Page 305 of this module, we assumed that the physical trainer was unaware of the underlying formulas and just used a website to get the data. Then from only three data points, we computed some equations. The official definitions are as follows

For someone who is x years old:

$$\begin{aligned} \text{Maximum Heart Rate} &= 220 - x \\ \text{Target Heart Rate, Upper Bound} &= 0.85(\text{Maximum Heart Rate}) \\ \text{Target Heart Rate, Lower Bound} &= 0.70(\text{Maximum Heart Rate}) \end{aligned}$$

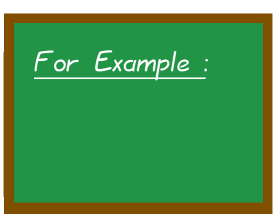
What makes this relevant is that we discovered—from a mere 3 data points (and 2 data points would have been enough)—our own equations which give almost the same answers as these official equations do. Why don't you test this by plugging your own age into these equations as well as the ones we found earlier, and seeing that the answers are very nearly the same—usually off by a fraction of 1 heartbeat. Now you can see the sheer power of the tools that we are developing.

Last but not least, you can see the significance of the slopes being 0.70 and 0.85. This defines "the zone," namely 70% to 85% of the maximum.

There is one other way to build a linear model, but I don't recommended it. The following equation is called the "point-point form" of a line:

$$y = \frac{y_1 - y_2}{x_1 - x_2}(x - x_1) + y_1$$

I dislike this formula because it is hard to memorize and memorization is accident prone. There are just too many operations, and it is easy to make an error. However, I have been told that this formula is popular among biologists. That does seem to make sense, because biologists are accustomed to memorization from courses like *Organic Chemistry*, *Anatomy and Physiology* as well as *Botany*.

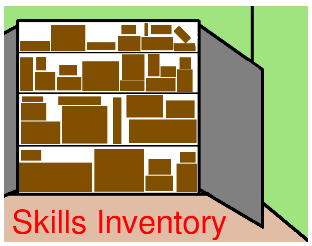


2-6-49

Let's see how to use the point-point form to produce a linear model for enrollment, using the points $(2011, 9357)$ and $(2010, 9339)$. We solved this earlier (on Page 299 of this module) quite easily, but with a few steps. The point-point form it is shorter, but more numerical.

$$\begin{aligned} y &= \frac{y_1 - y_2}{x_1 - x_2}(x - x_1) + y_1 \\ &= \frac{9357 - 9339}{2011 - 2010}(x - 2011) + 9357 \\ &= \frac{18}{1}(x - 2011) + 9357 \\ y &= 18(x - 2011) + 9357 \\ &\quad \ldots\text{or if } y = mx + b \text{ form is desired}\ldots \\ y &= 18x + 18(-2011) + 9357 \\ y &= 18x - 26,841 \end{aligned}$$

We will not have use of this formula again, and I recommend against memorizing it, because it is too easy to get it slightly wrong.



We have learned the following skills in this module:

- Finding the equation of a line through a given point with a given slope.
- Finding the equation of a line through two points.
- Constructing a linear model of a phenomenon with two data points.
- Checking the correctness of a model by plugging in some data points.
- Given three data points, constructing a model with the outer two and checking the model with the inner one.
- Evaluating the pros and cons of either point-slope form or slope-intercept form.
- Recognizing the boundaries of validity of a model, particularly when prices or the demand become negative numbers.
- Calculating the target heart rate, based upon age, for exercise.
- Representing error in a model in terms of residual error and relative error.
- Predicting, from a linear model, simple forecasts—while recognizing the inherent unreliability of such forecasts.
- Identifying the meaning of the slope of the line in several particular linear models.
- Realizing that the slope concept cannot be used to represent how a cubic function changes. Something more than a mere number is required.
- Constructing a linear model by solving a system of linear equations—the "Chinese method."
- Turning a linear model inside-out, which we will later rename "computing an inverse function."
- Computing the equation of a linear model using point-point form.
- As well as the vocabulary terms: domain, economies of scale, elastic goods, elasticity of demand, inelastic goods, point-slope form, saturation point, and slope-intercept form.

Here are the answers for the checkerboard box on Page 319 of this module asking about the slopes of the lines in the subscribers problem from Page 307 of this module.

- The slope for $w(t)$ was the change in the number of web-based subscribers per year. (The slope was positive, because that was increasing.)
- Meanwhile, the slope for $p(t)$ was the change in the number of print-based subscribers per year. (The slope was negative, because that was decreasing.)
- Finally, the slope for $b(t)$ was the change in the number of both types of subscribers per year. (The slope was positive, because that was increasing.)