Versatile Mathematics

COMMON MATHEMATICAL APPLICATIONS

Josiah Hartley
Frederick Community College

Val Lochman
Frederick Community College

Erum Marfani
Frederick Community College

2nd Edition

2020

Attributions This book benefited tremendously from others who went before and freely shared their creative work. The following is a short list of those whom we have to thank for their work and their generosity in contributing to the free and open sharing of knowledge.

- David Lippman, author of *Math in Society*. This book uses sections derived from his chapters on Finance, Growth Models, and Statistics. He also administers MyOpenMath, the free online homework portal to which the problems in this text were added.
- The developers of onlinestatbook.com.
- OpenStax College (their book *Introductory Statistics* was used as a reference)
  OpenStax College, *Introductory Statistics*. OpenStax College. 19 September 2013. <http://cnx.org/content/col11562/latest/>
- The authors of OpenIntro Statistics, which was also used as a reference.
- The Saylor Foundation Statistics Textbook: http://www.saylor.org/site/textbooks/Introductory%20Statistics.pdf

Thanks The following is a short list of those whom we wish to thank for their help and support with this project.

- The President's office at Frederick Community College, for providing a grant to write the first chapters.
- Gary Hull, who in his tenure as department chair gave us his full support and gave us the impetus to start the project, and generously shared his notes for MA 103.
- The entire FCC math department, who provided untold support and encouragement, as well as aid in reviewing and editing the text.

Contents

Chapter

1 Growth Models

What does the future hold? Although we cannot peer forward and find out for sure, it turns out that in many cases we can use mathematical models to make predictions. For instance, by tracking the development and path of a hurricane, we can predict where it will go next, although weather prediction is such a difficult problem that the true path is often very different from the prediction. While we won't be tackling anything nearly as complicated as weather forecasting, we will see in this chapter how to build a few different types of mathematical models and use them to make predictions.

Of course, it is crucial to understand that no mathematical model is perfect. There will always be a trade-off between the accuracy of a model and its simplicity. The simpler a model, the more easily we can make predictions with it, but there will be more error in the approximation. On the other hand, more precise models may be more difficult—or even impossible—to solve. You should always remember, though, that every model is at best an imperfect representation of the real world, and there will always be some inherent error between what the model predicts and what actually occurs.

SECTION 1.1 Linear Models

If you decide to train for a marathon and you currently run 3 miles a day, you may choose to increase your distance by half a mile every week. Can we predict how far you'll be running in six weeks, or how long it will take to reach your goal? With small numbers like these, you could answer those questions without using an equation, but we'll use this as an example to show how to build a simple model from a scenario like this.

Let P_t represent the number of miles that you run after t weeks, so P_0 would be the number of miles you currently run, P_1 would represent the number of miles you run after 1 week, and so on. We can define a **recursive relationship** like the following one to represent the scenario that was laid out.

$$
\begin{aligned}
P_0 &= 3 \\
P_t &= P_{t-1} + 0.5
\end{aligned}
$$

A recursive relationship is one that defines each value in a sequence using previous values. We could use this relationship to go from P_0 to P_1 to P_2 and so on, all the way to P_6 to answer the first question, and we could keep going from one value to the next until we reached 26 to answer the second question. However, this would be tedious and mindless, so instead we prefer an **explicit equation**, or closed-form equation. Especially for predictions far into the future, the recursive form is impractical, even though it arises easily from the problem description.

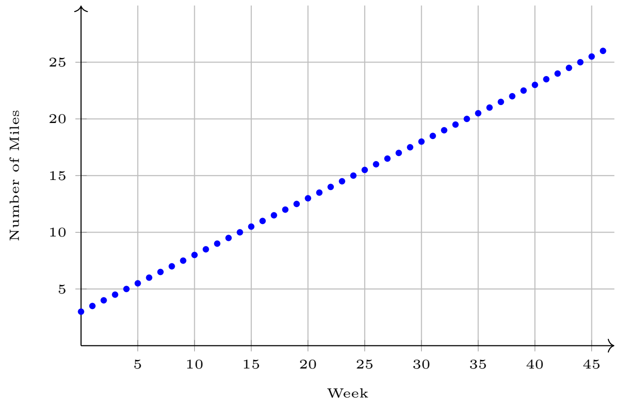
In this case, the explicit form is pretty straightforward, but deriving it from the recursive form will be instructive:

$$
\begin{aligned}
P_0 &= 3 \\
P_1 &= 3 + 0.5 \\
P_2 &= 3 + 0.5 + 0.5 = 3 + (0.5)2 \\
P_3 &= 3 + 0.5 + 0.5 + 0.5 = 3 + (0.5)3 \\
&\vdots \\
P_t &= 3 + 0.5t
\end{aligned}
$$

This explicit equation gives an easy way to quickly answer both questions. We can substitute 6 for t to find that $P_6 = 3 + 0.5(6) = 6$ miles, and we can substitute 26 for P_t and solve for t to find how long it'll take to reach your goal:

$$
\begin{aligned}
26 &= 3 + 0.5t \\
23 &= 0.5t \\
46 &= t
\end{aligned}
$$

It'll take 46 weeks to reach your goal.



This graph shows why we call this *linear* growth. If we graph the number of miles versus the week, the points lie along a straight line. This is consistent for every problem where a number grows by a constant amount every time period. That's the key to linear growth: there's a constant growth amount, which we'll call d (for *difference*), that is added at each step. For instance, in the marathon example, $d = 0.5$ because half a mile is added each week. Knowing that, the following formula should look familiar, because it is the same form that we developed above.

Linear Growth

If some quantity starts at size P_0 and grows by d every time period, then the quantity after t time periods can be predicted using the following formula.

$$P_t = P_0 + dt$$

Here d represents the common difference—the amount that the quantity changes each time t increases by 1.

Notice that this could refer to linear growth or linear decay; if d is negative, the quantity will decrease linearly.

Knowing that the key to linear growth is this common difference between terms, we can recognize linear growth from data if each term is the previous term plus a constant.

Term	Quantity	Difference from Previous Term
0	15	
1	27	12
2	39	12
3	51	12
4	63	12
5	75	12

As we can observe in this table, if we note that the quantity adds a constant amount each time, we know that the growth is linear, and we can write the closed-form equation given above.

Notice that this is exactly the standard linear equation that you've seen in your algebra classes:

$$y = mx + b$$
$$P_t = dt + P_0$$

Here, P_0 is the y-intercept, since it is the starting point, and thus the value when $t = 0$. Also, d is the slope here, or the amount by which the quantity changes when t increases by 1.

These two equations are the same, but as you'll see when modeling, we often rename the pieces to more closely match the names of the real-world quantities we're measuring.

EXAMPLE 1 ELK POPULATION

The population of elk in a national forest was measured to be 12,000 in 2011 and 15,000 in 2015. If the population continues to grow linearly at this rate, what do we expect the elk population to be in 2022?

We first need to define the parts of our linear growth equation. The initial amount P_0 is the amount when $t = 0$, but we won't use the actual year 0 as our starting point. Instead, the initial amount in this problem is given in 2011, so we'll define $t = 0$ to be the year 2011, so $P_0 = 12,000$.

Next we need to find d, the growth per time period. Since the time period in this example is one year, we'll need to find how much the population grew each year.

Year	Population
0	12,000
4	15,000

Since the population grew by 3,000 in 4 years, this represents a growth of 3,000/4 = 750 per year. Thus $d = 750$.

Note that this is equivalent to using the slope formula: $\dfrac{\text{rise}}{\text{run}}$

$$\begin{aligned} d &= \text{slope} \\ &= \frac{\text{change in population}}{\text{change in time}} \\ &= \frac{15,000 - 12,000}{2015 - 2011} \\ &= \frac{3000}{4} \\ &= 750 \end{aligned}$$

Now we can write the explicit equation that models this population growth:

$$P_t = 12,000 + 750t$$

To answer the question, we note that 2022 corresponds to $t = 11$, since 2022 is 11 years after 2011.

$$\begin{aligned} P_{11} &= 12,000 + 750(11) \\ &= \boxed{20,250 \text{ elk}} \end{aligned}$$

TRY IT

If we estimated the population of trout in a pond to be 2200 in 2008 and 3500 in 2012, construct a linear model to predict the population in 2017.

Notice the key to that last problem: by knowing the population at the beginning and at the end, we can how much the population changes in a single year by dividing the total change by the number of years that elapsed.

In the next example, we'll look at data that is nearly linear, but not exactly. However, for the purposes of making predictions, we'll treat it as if it follows a linear trend (remember, in the real world, data is messy). We'll do exactly the same thing we just did in the previous example; we'll simply focus on the amount at the beginning and at the end, and divide the total growth by the amount of time that passed.

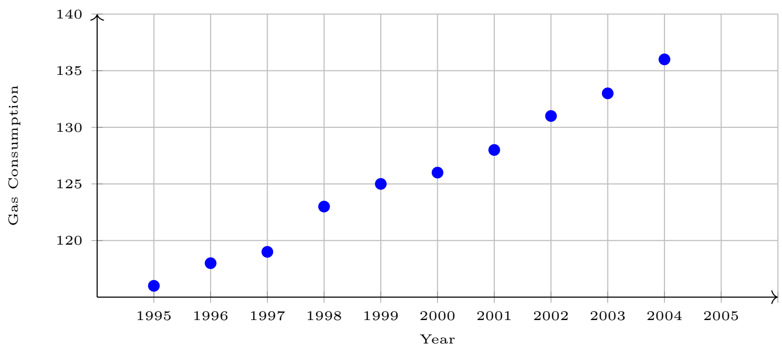
GASOLINE CONSUMPTION

EXAMPLE 2

Gasoline consumption in the US has been increasing steadily. Data from 1995 to 2004 is shown below. Find a linear model for this data, and use it to predict consumption in 2018. If the trend continues, when will consumption reach 200 billion gallons?

Year	’95	’96	’97	’98	’99	’00	’01	’02	’03	’04
Consumption (billions of gallons)	116	118	119	123	125	126	128	131	133	136

If we plot this data, it appears to have an approximately linear relationship.



One way to find a linear trend for data like this is a statistical technique known as *linear regression*. We will see examples of how to use the calculator for regression later in this section, and the chapter on Statistics will describe linear regression in more detail.

For now, though, we’ll simply use the data from the first and last years to find the average growth each year (the slope of the line).

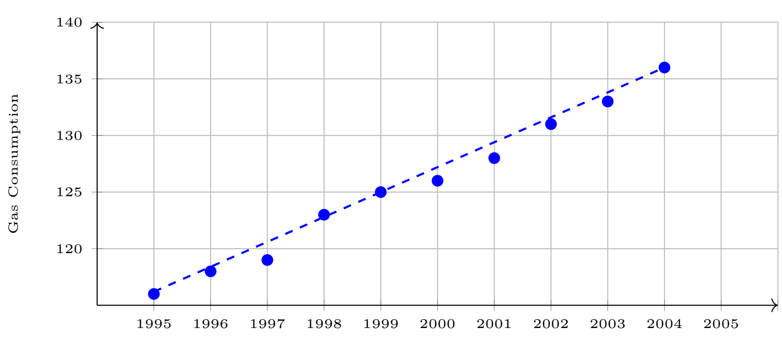
We could use the data from any two years to calculate the slope (and we would get slightly different answers), but a common convention is to use the first and last years. You should follow this convention when answering the homework questions.

Year	**Consumption**
1995	116
2004	136

$$d = \text{ slope } = \frac{\text{change in consumption}}{\text{change in time}} = \frac{136 - 116}{2004 - 1995} = \frac{20}{9}$$
$$= 2.22 \text{ billion gallons per year}$$

Now we can write our model (in billions of gallons):

$$\boxed{P_t = 116 + 2.2t}$$



We can use our model to make predictions about the future, using the simplifying assumption that the previous trend continues unchanged.

This example illustrates the two main types of questions that we often want to answer:

1. Predicting the value of what we are measuring at a given point in time.
2. Predicting the point in time when the thing we are measuring will reach a certain value.

- Predicting gas consumption in 2018, when $t = 23$:

$$P_{23} = 116 + 2.2(23) = \boxed{166.6}$$

Our model predicts that the US will consume 166.6 billion gallons of gasoline in 2018 if the current trend continues.

- Predicting when consumption reaches 200 billion gallons:

$$\begin{aligned} 200 &= 116 + 2.2t \\ 84 &= 2.2t \\ \boxed{38.18} &= t \end{aligned}$$

This model predicts that gas consumption will reach 200 billion gallons about 38 years after 1995, or the year 2033.

TRY IT

The number of stay-at-home fathers in Canada has been growing steadily at an approximately linear rate. Use the data from the table below to find an explicit formula for the number of stay-at-home fathers and use it to predict the number in 2020. Use 1976 and 2010 to find the average rate of change.

Year	1976	1984	1991	2000	2010
Number of stay-at-home fathers	20,610	28,725	43,530	47,665	53,555

Again, we understand that this model is not perfect; the US will most likely not consume exactly 166.6 billion gallons of gas in 2018, but we expect consumption to be *about* that. In practice, we'll often make predictions and then compare them to actual measured results to assess the accuracy of our model. A very simple linear model like this will likely have fairly large error; more sophisticated models tend to have smaller errors.

Predicting Time

As we pointed out in the previous example, there are generally two questions that we'll encounter with mathematical models like the ones in this chapter:

- At a given time, find the amount
- For a given amount, find the time it will take to reach that level

The first question is generally easier, simply because of how the formula is arranged:

$$P_t = P_0 + dt$$

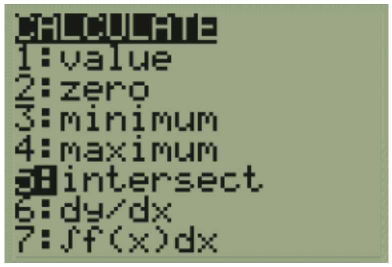
Since we know everything on the right-hand side of the equation in this case, we can plug it all in and simply carry out the arithmetic; what we want to know is already isolated on the left side.

It's slightly more tricky to solve for t in the second question, because it requires some algebra to rearrange the pieces of the formula so that t will be isolated on one side. This gets even more difficult for more complicated models, like the ones in the rest of the chapter, because the algebraic steps are more involved.

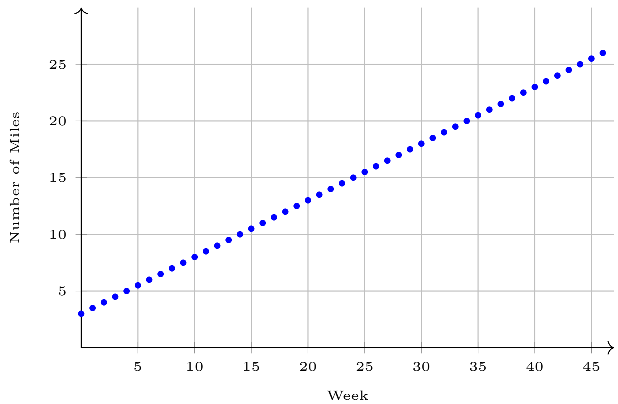
Another option: use a calculator

Although you should be able to carry out the algebraic steps, and it will likely be quicker to do so with linear models, let's see how to use the calculator to accomplish the same purpose. What we're about to see will work no matter which part of the equation we want to solve for, so we will return to this method in each section in this chapter whenever we want to solve for t.

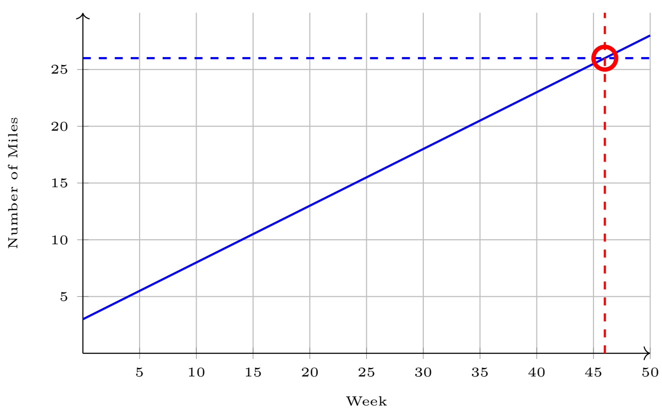
The key to this method involves graphing both sides of an equation and using the `INTERSECT` option on a graphing calculator. If you have a TI-83 or 84 or similar, you can look under the `CALC` menu, which you can find by pressing 2ND and TRACE:



Before diving into the details, let's see why this works by going all the way back to the opening example of this section: the marathon training program. Remember that we drew this graph to represent the number of miles that you would run each week:



Now, let's ask this question: how long will it take to reach the goal of running 26 miles? Notice that what this means on the graph is: when does that blue line reach a height of 26? If we draw a horizontal line at 26, we can zero in on the point where the two lines cross (or intersect) and trace downward to the x-axis to find the time when that intersection occurs.



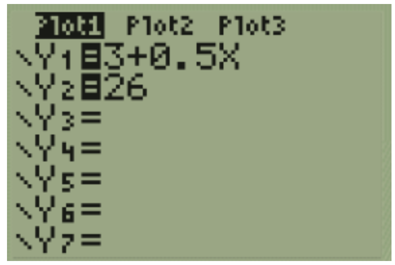
In other words, when we find the coordinates of the intersection point, the y-coordinate will be the level we're given to predict the time for, and the x-coordinate will be the time at which that level is reached.

On the calculator, this means we have to graph two functions: one will be the model that we have built ($P_t = P_0 + dt$ for linear models in this section, other formulas for later sections), and the other will be a constant value, whatever value we want to know the corresponding time for. Once we find the intersection, the x-coordinate will be the time value that we're looking for.

To illustrate how to use this method, we'll continue to use the same example: remember that the model we built was $P_t = 3 + 0.5t$, and we wanted to predict when the number of miles would reach 26. Thus, we need to graph $y_1 = 3 + 0.5x$ and $y_2 = 26$ on the calculator; note that the calculator always uses x and y instead of t and P_t.

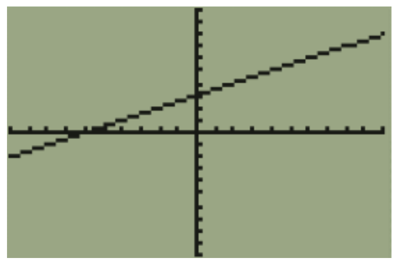
Step 1: graph both functions

To begin, press the $\boxed{\text{Y=}}$ button in the upper-left corner to open the graphing menu, then enter the two functions as Y_1 and Y_2. Note that to enter X, use the button labeled $\boxed{X, T, \theta, n}$



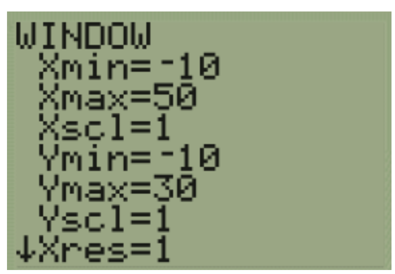
Step 2: zoom to see the intersection

When you press $\boxed{\text{GRAPH}}$ you may see something like this:

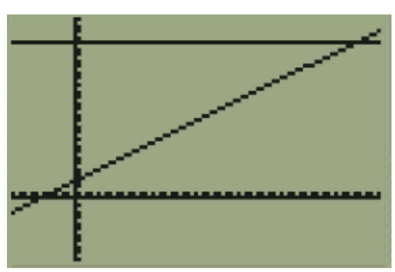


Only one line is visible, because the window is too zoomed in to see the other. To zoom out, you can press the $\boxed{\text{ZOOM}}$ button and select one of the options, like `3. Zoom Out`. If you press the $\boxed{\text{WINDOW}}$ button, though, you can manually set the boundaries of the window, and if you have a sense of roughly where the intersection should occur, that's usually the easiest way to adjust the view so that you can see both lines.

In this case, we know that the upper end of the window needs to go up to at least 26, since that is the y-coordinate of the intersection, so let's use 30 as the upper limit. With a bit of trial and error, we can find a reasonable upper limit on x as well; we'll use 50 for this one.



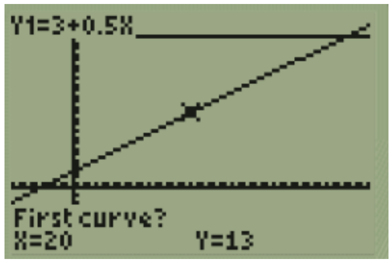
If you press $\boxed{\text{GRAPH}}$ again, you should see this:



We can now see the intersection point on the graph, but we have a bit more work to do to find its coordinates.

Step 3: find the intersection

Now if you press 2ND and TRACE to open the `CALC` menu and select the option labeled `5. intersect`, you'll see something like this:



What's happening here? Without getting too specific, the way the calculator finds the intersection point involves starting with an initial guess, and then searching nearby in both directions to find the crossing point. This means it will ask where you want to start looking on the first curve, where you want to start looking on the second curve, and where your overall initial guess is. You can use the arrow keys to move the blinking cursor closer to the intersection if you like, but for simple examples like the ones in this chapter, the calculator will find it even if we don't feed it a starting point that is close by.

In practice, this means that if you like, you can simple press ENTER three times to accept the default starting values.

Once you do this, you should see the following:

At the bottom of the screen, you can see the coordinates of the intersection point. Notice that the y-coordinate is no surprise; that's based on the given number of miles that we were trying to reach. The x-coordinate is the answer we were looking for, the value of t when we reach 26 miles.

Linear Regression Using the Calculator

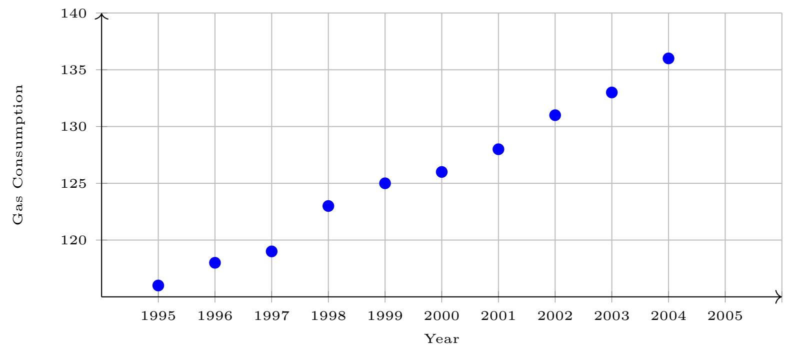
Usually when we build a model, we do so using data from the past to predict the future. In most of the examples in this section, we drew a straight line by selecting the first and last points and finding the slope that connected them.

There is another approach, using a statistical technique known as **linear regression**. We will omit the details here, but a more thorough discussion can be found in the Statistics chapter. For our purposes, we will simply see how to use a graphing calculator to find the equation of a line that most closely fits the data we're given.

Rather than simply using two points, linear regression takes all of the data points into account and builds a *line of best fit*; the principles needed to describe how this best-fit process works require some statistics knowledge, so we will not discuss it further in this chapter.

We'll use the gasoline consumption example to show how to use the calculator for this. Recall that we were given the following data (and the resulting graph):

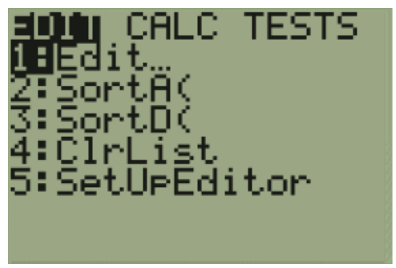
Year	'95	'96	'97	'98	'99	'00	'01	'02	'03	'04
Consumption (billions of gallons)	116	118	119	123	125	126	128	131	133	136



The model we built in that example was $P_t = 116 + 2.2t$; we'll keep this handy to compare it to the answer we get using regression.

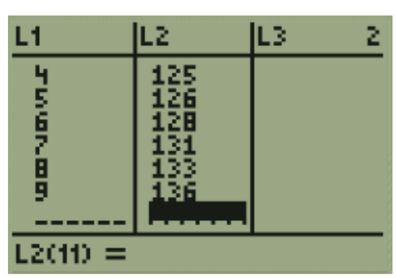
STEP 1: ENTER THE DATA

First, we need to enter the data that we're given (in full). To do this, press the $\boxed{\text{STAT}}$ button; the first menu that you will see should look like this:



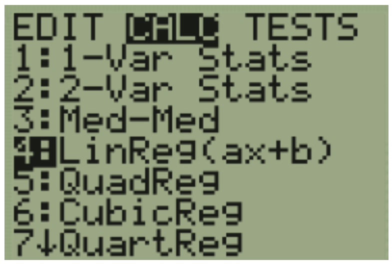
Simply press $\boxed{\text{ENTER}}$ to enter the `Edit` menu, where we can enter the data; first, you should see the following (if there is data already there, simply use the arrow keys to move up to the label of one of the lists, and press $\boxed{\text{CLEAR}}$):

In list `L1`, enter the values for t. Notice that here we set 1995 as the initial year, or year 0, so start with 0 and enter enough values to get to 2004 (year 9). In list `L2`, enter the values for gasoline consumption.

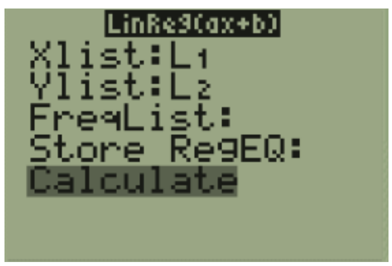


STEP 2: CALCULATE THE REGRESSION EQUATION

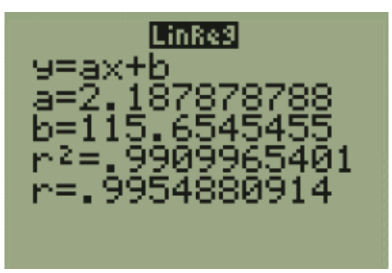
Now, to calculate the regression equation, press the STAT button again, and use the right arrow key to move to the CALC menu. There are many options here, but the one we want at the moment is 4: LinReg(ax+b). Notice that this means that we'll receive answers for a and b, and these will represent the slope and intercept, respectively.



If you select the linear regression option, you should see the following menu:



Since we entered the time values in L1 and the consumption values in L2, we don't need to change anything here, so simply scroll down to select Calculate and press ENTER, which shows the results:



For now, all we're interested in is the values of a and b; to put it in more familiar terms, the equation of the regression model is

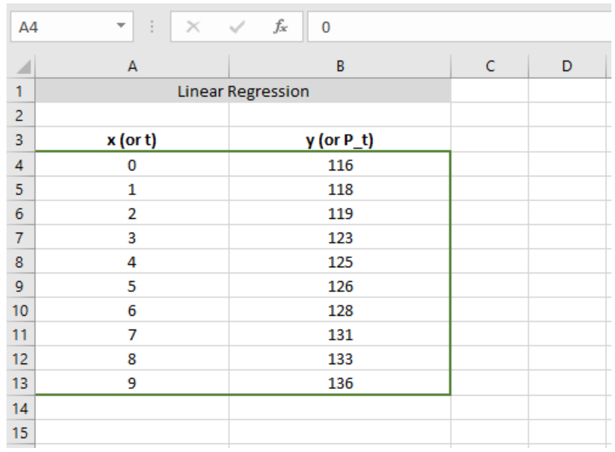
$$P_t = 115.7 + 2.19t$$

Compare this to the one we built by hand. Since this data was fairly linear, there isn't much difference between the two answers; if there were *outliers* (unusual values that deviate significantly from the linear trend), there might be a greater difference.

In general, the manual method is easier and quicker to do, but this statistical method gives more accurate results.

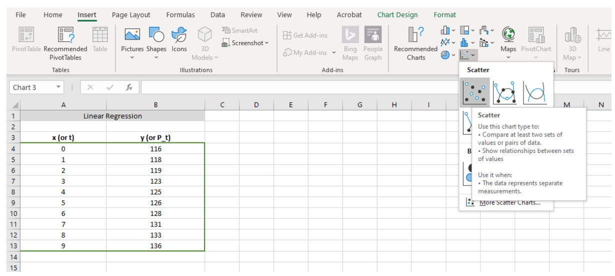
Using Excel

We can also use Excel to calculate regression equations. To begin, enter the data as shown (we're using the same example as we did with the calculator, for comparison):

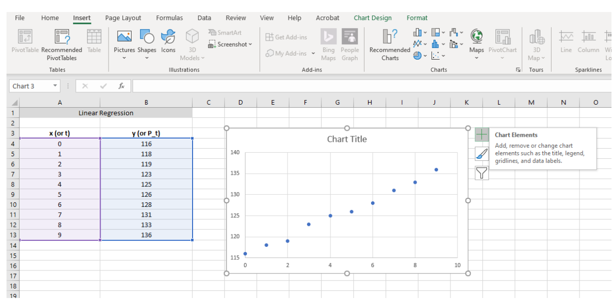


	A	B
1	Linear Regression	
2		
3	x (or t)	y (or P_t)
4	0	116
5	1	118
6	2	119
7	3	123
8	4	125
9	5	126
10	6	128
11	7	131
12	8	133
13	9	136

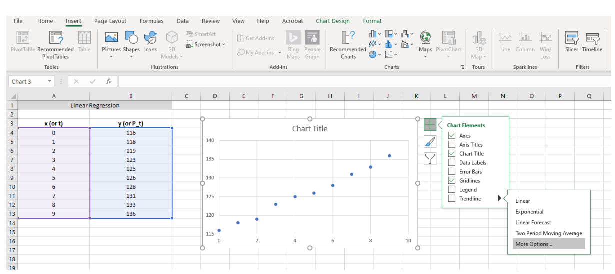
Next, we need to insert a scatterplot of these data points; select the Insert menu at the top of the screen, then select the scatterplot option under the Charts section (select the first type of scatterplot under that submenu).



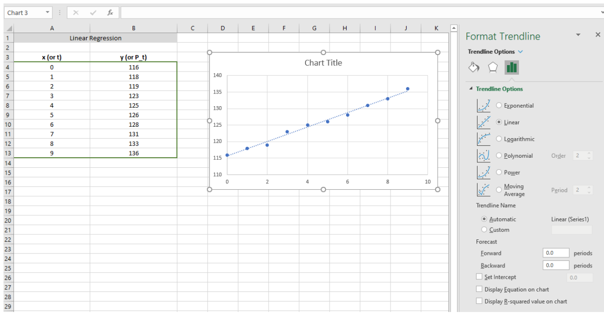
This creates a graph similar to the one that we drew earlier for this same example; we could add a title to the chart and the axes, but we won't bother for this example.



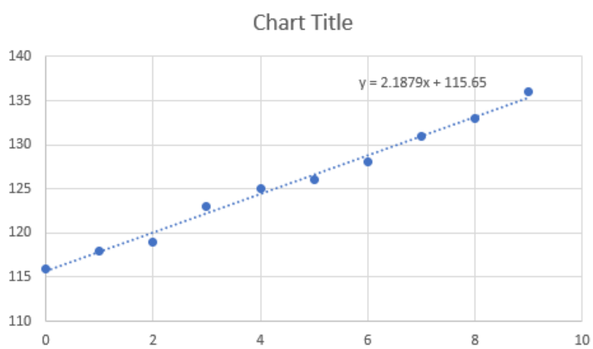
To add a regression line, click on the plus symbol at the upper right (when the chart is selected). One of the options is to add a trendline. If you check that box, the line will be drawn, but by default, Excel won't show the equation of the line, which is what we're after, so we need to select "More Options" after clicking on the arrow next to the trendline option.



This brings up the following menu. Notice that there are different types of equations we can use to model the data, including several that we'll encounter later in the chapter. For now, though, leave the linear option selected, but check the box at the bottom of the menu that says "Display Equation on chart."



This leads to the following:



The equation, as calculated by Excel, is

$$P_t = 115.65 + 2.1879t$$

which, after rounding, is identical to the one obtained by the calculator (this is no surprise, since they're using the same formulas).

When Good Models Go Bad

When predicting the future with mathematical models, it is crucial to keep in mind that few trends continue indefinitely.

A BOY'S HEIGHT — EXAMPLE 3

Suppose a four year old boy is currently 39 inches tall, and you are told to expect him to grow 2.5 inches a year.

We can set up a growth model, with $t = 0$ corresponding to 4 years old.

$$P_t = 39 + 2.5t$$

At six years old (when $t = 2$), we would expect him to be

$$P_2 = 44 \text{ inches tall,}$$

but this model eventually breaks down. Certainly, we shouldn't expect him to grow at the same rate all his life. If he did, at age 50 he would be

$$P_{46} = 154 \text{ inches } = 12.8 \text{ feet tall.}$$

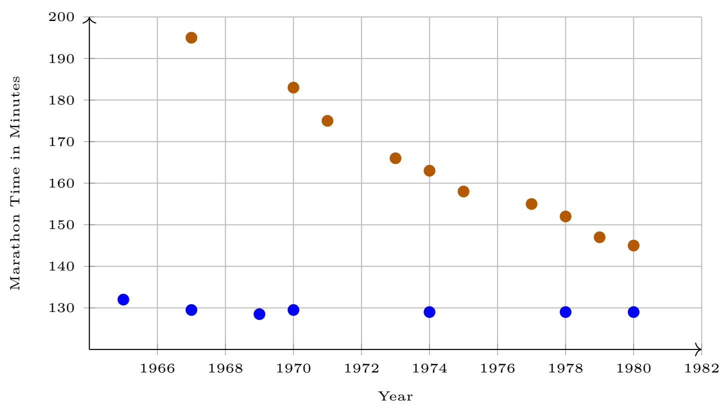
Of course, this boy will not grow at a constant rate, but rather experience growth spurts and ultimately stop growing in his early 20s. But this example also illustrates that we should check our model against common sense.

Let's look at another example that illustrates the need for a common sense check.

EXAMPLE 4 MARATHON TIMES

The table and graph below show the record times for the marathon for men and women from 1965 to 1980.

Year	Men's Times (min)	Women's Times (min)
1965	132	
1966		
1967	129.5	195
1968		
1969	128.5	
1970	129.5	183
1971		175
1972		
1973		166
1974	129	163
1975		158
1976		
1977		155
1978	129	152
1979		147
1980	129	145



From this data, it looks like both sets of data are following a linear trend. If we use the first and last data points to find the average rate of change for each, we get the following linear models, using 1967 as $t = 0$:

$$M_t = 129.5 - 0.2t$$
$$W_t = 195 - 3.85t$$

According to these two linear models, we would predict that the women's record would beat the men's record by 1985; however, in 1985, the men's record was still 14 minutes faster than the women's. What happened here?

Since women began setting marathon records about 50 years later than men, in the early years their progress was drastic, but eventually slowed down, and the trend was not linear over the long run (wow, what a terrible pun).

It should be clear that this linear trend was misleading, since if we extrapolated this model too far forward, we'd get ridiculous results. The model predicts, for instance, that women would run the marathon in 1:20:00 in 1997 (a pace of about 20 mph, the speed of a roadrunner or close to the top speed of Usain Bolt at full sprint), or that by 2017 they'd be running it in 2.5 minutes (around 630 mph).

The lesson is simple, and hopefully obvious: linear trends are usually only useful in the short term; few phenomena follow linear trends over the long term. That is why we'll examine other types of models in the coming sections. However, keep this in mind, because we'll find that even those more sophisticated models have their limitations, and often they too break down in the long term.

Exercises 1.1

1. Marko currently has 20 tulips in his yard. Each year he plants 5 more.

(a) Find a linear model of the form $P_t = P_0 + dt$ to describe the number of tulips Marko has at a given point in time.

(b) How many tulips will Marko have in 7 years?

(c) When will Marko have 65 tulips?

2. Pam is a DJ. Every week she buys 3 new albums to add to her collection. She currently owns 450 albums.

(a) Find a linear model of the form $P_t = P_0 + dt$ to describe the number of albums Pam has at a given point in time.

(b) How many albums will Pam have in 11 weeks?

(c) When will Pam have 489 albums?

3. A store did $40,000 in sales in 2016, and $62,000 in 2018.

(a) Assuming the store's sales are growing linearly, find the growth rate d.

(b) Write a linear model to describe this store's sales from 2016 onward.

(c) Predict the store's sales in 2025.

(d) When do you expect the store's sales to exceed $100,000?

4. A small town had 340 homes in 2010, and by 2020, this had grown to 375.

(a) Assuming the number of homes is growing linearly, find the growth rate d.

(b) Write a linear model to describe the number of homes in this town from 2010 onward.

(c) Predict how many homes there will be in this town in 2035.

(d) When do you expect the number of homes to exceed 500?

5. A population of beetles is growing according to a linear growth model. Initially, there were 13 beetles, and 8 weeks later, there were 42 beetles.

(a) Write a linear model to describe the number of beetles over time, using weeks as the unit of time.

(b) How many beetles are there expected to be 14 weeks after the initial point?

(c) When do you expect the number of beetles to exceed 70?

6. The number of streetlights in a town is growing linearly. Four months ago there were 130 lights. Now there are 146 lights.

(a) Write a linear model to describe the number of streetlights in the town over time, using months as the unit of time.

(b) How many streetlights are expected in a year?

(c) When do you expect the number of streetlights to exceed 200?

7. In 1990, there were 112 nuclear power plants in the U.S. By 2019, this number had fallen to 96.

(a) Write a linear model to describe the number of power plants from 1990 onward.

(b) Using this linear model, predict the number of power plants in the U.S. in 2030.

(c) When do you expect the number of power plants to reach 90?

8. In 1990, approximately 1,820,000 violent crimes were reported in the U.S. By 2019, this number had fallen to approximately 1,200,000.

(a) Write a linear model to describe the number of violent crimes in the U.S. from 1990 onward.

(b) Using this linear model, predict the number of violent crimes in 2040.

(c) When do you expect the number of violent crimes to reach 1,000,000?

9. The table below shows the average annual cost of health insurance for a single individual, from 1999 to 2019, according to the Kaiser Family Foundation.

Year	**Cost**
1999	\$2,196
2000	\$2,471
2001	\$2,689
2002	\$3,083
2003	\$3,383
2004	\$3,695
2005	\$4,024
2006	\$4,242
2007	\$4,479
2008	\$4,704
2009	\$4,824
2010	\$5,049
2011	\$5,429
2012	\$5,615
2013	\$5,884
2014	\$6,025
2015	\$6,251
2016	\$6,435
2017	\$6,690
2018	\$6,896
2019	\$7,186

(a) Using only the data from the first and last years, build a linear model to describe the cost of individual health insurance from 1999 onward.

(b) Using this linear model, predict the cost of insurance in 2030.

(c) According to this model, when do you expect the cost of individual insurance to reach \$10,000?

(d) Using a calculator or spreadsheet program, build a linear regression model to describe the cost of individual insurance from 1999 onward.

(e) Using the regression model, predict the cost of insurance in 2030.

(f) According to the regression model, when do you expect the cost of individual insurance to reach \$10,000?

10. The table below shows the average annual cost of health insurance for a family, from 1999 to 2019, according to the Kaiser Family Foundation.

Year	**Cost**
1999	\$5,791
2000	\$6,438
2001	\$7,061
2002	\$8,003
2003	\$9,068
2004	\$9,950
2005	\$10,880
2006	\$11,480
2007	\$12,106
2008	\$12,680
2009	\$13,375
2010	\$13,770
2011	\$15,073
2012	\$15,745
2013	\$16,351
2014	\$16,834
2015	\$17,545
2016	\$18,142
2017	\$18,764
2018	\$19,616
2019	\$20,576

(a) Using only the data from the first and last years, build a linear model to describe the cost of family health insurance from 1999 onward.

(b) Using this linear model, predict the cost of insurance in 2025.

(c) According to this model, when do you expect the cost of family insurance to reach \$30,000?

(d) Using a calculator or spreadsheet program, build a linear regression model to describe the cost of family insurance from 1999 onward.

(e) Using the regression model, predict the cost of insurance in 2025.

(f) According to the regression model, when do you expect the cost of family insurance to reach \$30,000?

SECTION 1.2 Quadratic Models

In reality, straight lines are rare; this is as true with data as it is with stone formations. Real data hardly ever follows such a simple rule, but whenever we can, we like to use lines to approximate real trends because straight lines are easy to work with.

However, as we found at the end of the previous section, linear models will not work in every situation, and trying to force a linear trend on real data often leads to nonsensical results. For the rest of this chapter, we'll study other models–using slightly more complicated formulas–that can be used when the data does not follow a linear trend.

Picking which model to use is a hard question, and this decision is usually based on looking at a graph and looking for a pattern, or trying multiple models and seeing which gives the best results. That kind of decision is largely beyond the scope of this chapter; we'll focus on learning some basic features of a few different models, and the questions in each section will specify which kind of model to use.

We'll discuss three types of non-linear models:

1. Quadratic Models
2. Exponential Models
3. Logistic Models

This section will focus on quadratic models, and the last two types will be the topics of the remaining sections in this chapter.

Quadratic Models: Parabolas

The next time you use a water fountain, watch the path that the water follows. There's a certain elegance to this arch, and we have a special name for it: a **parabola**. Every time a baseball player throws a ball, that ball's path follows this same pattern. In fact, every time anything is launched or thrown, it follows a parabolic arch.

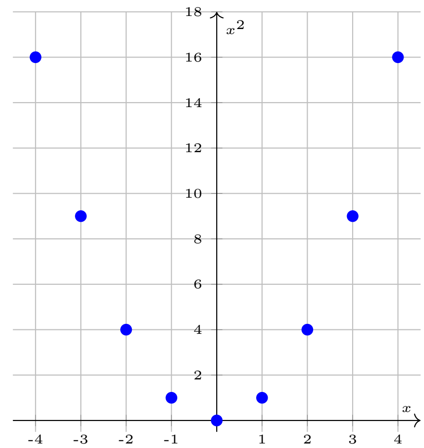
It turns out that this shape comes from a simple mathematical operation: *squaring*. Yes, that's it; when you square numbers, this pattern emerges.

Of course, that's not clear immediately, so let's do a simple experiment: square the whole numbers from 0 to 4, as well as the matching negative numbers. We can arrange our results in a table like this:

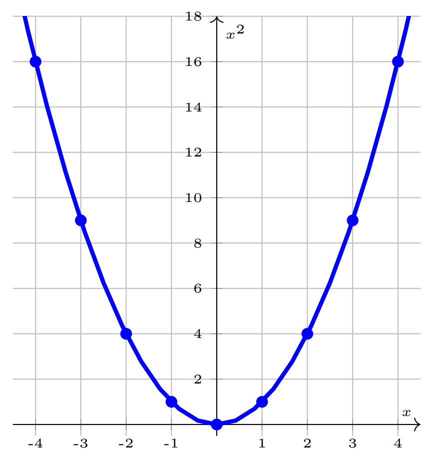
x	-4	-3	-2	-1	0	1	2	3	4
x^2	16	9	4	1	0	1	4	9	16

Notice the symmetry in the results: when we square 2 and -2, for instance, we get the same result, because multiplying two negative numbers results in a positive answer. Look at the picture of the fountain in the margin; one of the main features of a parabola is its symmetry, and we're already starting to see that emerge.

Just like we did in the previous section, let's graph these results and see what visual pattern we can observe.



And if we connect the dots by filling in the results we would get if we applied this same rule to non-whole numbers, what do we get?



Look at that: we got exactly the same shape as the water rising and falling in the fountain (although flipped upside-down). It's amazing what beauty can come from such a simple mathematical rule: when you start squaring numbers, you find a parabolic curve, and it turns out that this curve governs much of the motion we see around us every day.

Now, by comparing this graph with the picture of the fountain, we can observe that while parabolas all have the same structure, there are differences in the exact paths. For instance, some of the jets in the fountain are aimed higher than others, leading to taller, narrower arches, while the lower jets create shorter, wider arches. And the graph shows a parabola that is oriented in the opposite direction.

Where do these differences come from? It's important to remember that the basic structure of the parabola is generated by squaring the inputs (we used t in the previous section; these are the values for which we want to make predictions) to get x^2 (or t^2 if we use t as the label for our variable). A more general version of this function (and the formula we'll use when we start building quadratic models) is

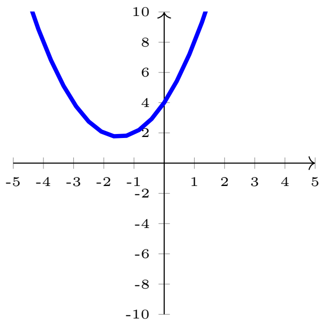
$$f(x) = ax^2 + bx + c$$

or, in a form that is more similar to what we used in the last section,

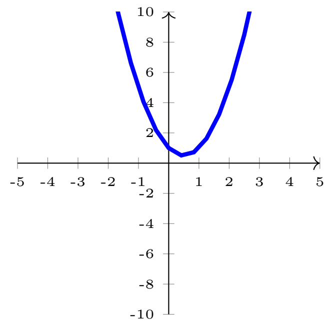
$$P_t = at^2 + bt + c.$$

Notice that if we set $a = 1$, $b = 0$, and $c = 0$, this simplifies to $f(x) = x^2$, which is the one we graphed above (the notation $f(x) =$ simply means that we're thinking of this as a function, a rule that takes inputs x and yields outputs, as we did in the table above).

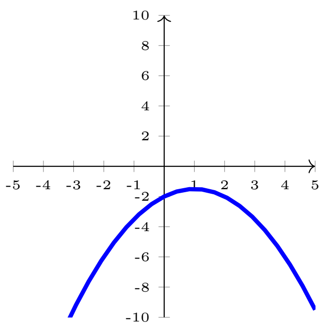
When we start changing the values of a, b, and c, that's when the variations among different parabolic shapes start to appear. Here are a few examples:



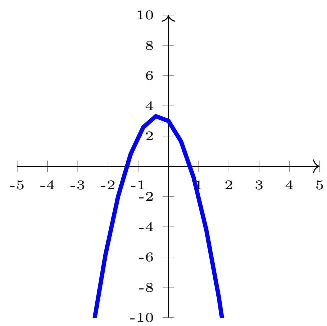
$a = 1,\ b = 3,\ c = 4$
$x^2 + 3x + 4$



$a = 2,\ b = -2,\ c = 1$
$2x^2 - 2x + 1$



$a = -0.5,\ b = 1,\ c = -2$
$-0.5x^2 + x - 2$



$a = -3,\ b = -2,\ c = 3$
$-3x^2 - 2x + 3$

There's no need to study this chart in much detail, but we can quickly gather a few observations by comparing the values of a, b, and c:

- When a changes, the shape of the curve changes. More specifically,
  - When a is positive, the curve opens upward; when a is negative, the curve opens downward.
  - Larger values of a (in magnitude) make the curve narrower (taller); smaller values of a make the curve wider (shorter).
- When b and c change, the location of the curve changes.
  - Increasing c moves the curve upward; decreasing c moves it downward.
  - The effect of b is harder to see, but it, in combination with a and c, moves the curve both vertically and horizontally.

The most important takeaway from this comparison is probably the effect of a. When we start fitting parabolas to data, the value of a, the number multiplied by x^2 or t^2, will tell us the direction of the curve, as well as something about how steep the curve is.

With this background, we're ready to state the formula that we can use to describe data with a quadratic model.

Quadratic Growth Models

The following formula can be used to make predictions when data follows an approximately parabolic trend:

$$P_t = at^2 + bt + c$$

The values of a, b, and c control the shape and location of the curve.

Notice that we use t in the formula, to match the formulas in the other sections in this chapter, but we'll use x interchangeably with t in the problems we do, since we'll use calculators and other technology to solve these problems, and those tools generally use x as the variable.

Using the Calculator

EXAMPLE 1 PLOTTING POINTS ON A CALCULATOR

According to the U.S. Census Bureau, the number of Americans over the age of 100 is increasing. The Census Bureau reported the following data, where the number of people is measured in the thousands:

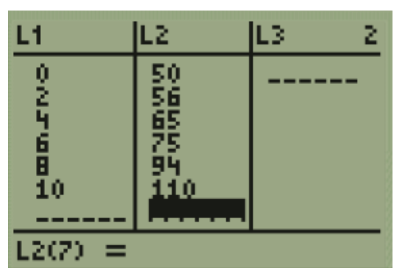
Year	Number (thousands)
1994	50
1996	56
1998	65
2000	75
2002	94
2004	110

Graph this data using a graphing calculator.

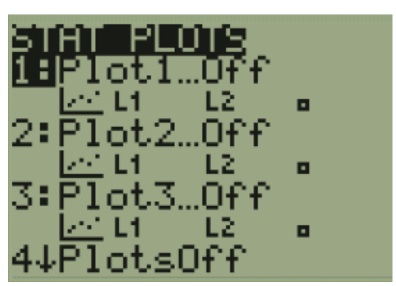
Solution To begin, we need to enter the data, using the same steps outlined in the previous section. Press the STAT button to enter the statistics menu, press ENTER to enter the data entry menu, and add the values in the table above to the table in the calculator. Use the same order in the columns.

Note: order of the columns This is discussed more in the Statistics chapter, but the first column generally refers to the variable that we use to predict the other. In this case, we can predict the number of Americans over the age of 100 by picking a given year. We would be less likely to predict what year we're describing based on the number of Americans over the age of 100.

We'll use 1994 as the beginning of our experiment, so that will correspond to year 0. Once you enter the data, you should see the following.

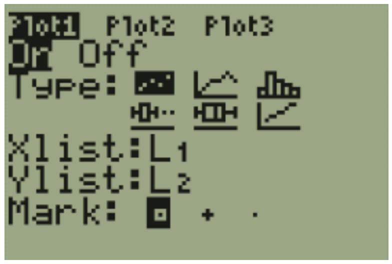


To graph this data, we need to use the `STAT PLOT` option, which is the `2ND` option on the Y= button, so to access it, press 2ND followed by Y= . This opens the following menu.

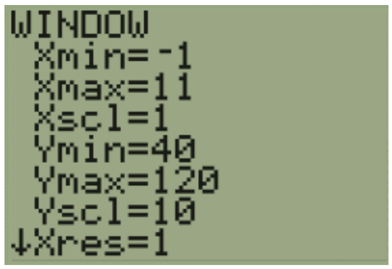


Pressing ENTER opens the menu for the first plot. In order to display a statistics plot, you must turn it on (and turn it off later if you'd like to only graph a function).

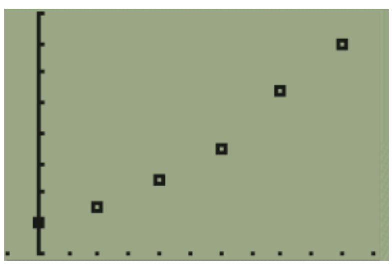
There are various types of statistics plots that the calculator can draw, but the first option (the one selected as shown) is a scatterplot, which is what we want. Leave the rest of the options as they are (note the order of the data columns).



We must zoom the window properly to ensure that we can see the data as it is graphed; the values shown below work well.



Now, if you press $\boxed{\text{GRAPH}}$ you should see the following:



Notice that the shape is similar to the parabolas we have seen drawn, or at least a portion of one.

TRY IT

Use a graphing calculator to plot the data below, and see if it follows an approximately quadratic trend.

x	y
0.9	2.5
1.3	4.1
1.6	5.1
2.1	7.5
2.5	9.8
3.2	14.3
3.6	18.1
4.2	23.0

The next question, of course, is how to fit a quadratic model to data like this. Specifically, this means finding values for a, b, and c that will make the resulting curve match the data as closely as possible.

Unlike with linear models, we won't do this manually, since the algebra necessary is more complicated than we'd like to tackle here. Thus, we will simply take advantage of the **quadratic regression** option on the calculator, which is labeled `QuadReg`.

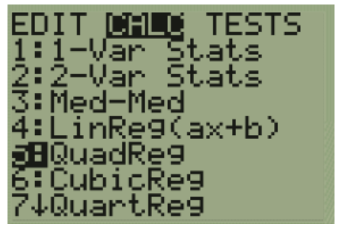
This can be found in the same menu as the linear regression function, and the process is very similar as well.

EXAMPLE 2 FITTING A QUADRATIC MODEL ON A CALCULATOR

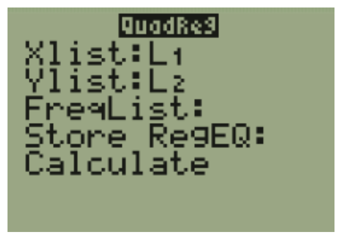
Using the same census data as in the previous example, find a quadratic model that can be used to predict how many Americans will be over the age of 100 in a given year.

Solution First, we must have the data entered, so if you don't have the data still in your calculator, follow the first step of the previous example to enter it.

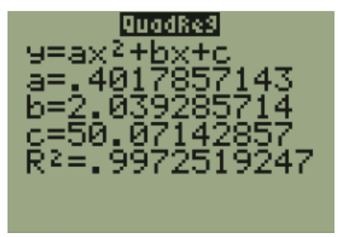
Next, press the STAT button and use the right arrow button to move to the `CALC` menu. There you should find an option labeled `5: QuadReg` that will perform quadratic regression.



If you entered the year in the first column and the number of people in the second column, you won't have to change anything on the next menu. Simply scroll down to `Calculate`, and press ENTER



The results are shown below.



Based on the values of a, b, and c given, we get the following model:

$$\boxed{P_t = 0.40t^2 + 2.04t + 50.07}$$

TRY IT

Use a graphing calculator to find a quadratic model for the data given below.

x	y
0.9	2.5
1.3	4.1
1.6	5.1
2.1	7.5
2.5	9.8
3.2	14.3
3.6	18.1
4.2	23.0

MAKING PREDICTIONS WITH A QUADRATIC MODEL

EXAMPLE 3

A study designed to track the gas mileage of a car based on its speed found the following results.

Speed (mph)	Mileage (mpg)
15	22.3
20	25.5
25	27.5
30	29.0
35	28.8
40	30.0
45	29.9
50	30.2
55	30.4
60	28.8
65	27.4
70	25.3
75	23.3

(a) Use a graphing calculator to plot the data.

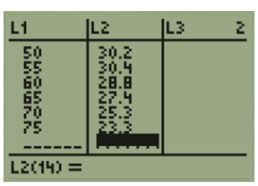
(b) Find a quadratic model that best fits the data.

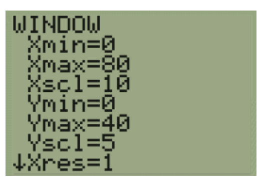
(c) Based on this model, what gas mileage should be expected at 62 miles per hour? At 90 miles per hour? Which of these predictions is likely to be more reliable?

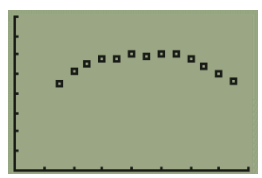
(d) Based on the model, what speeds are likely to produce a mileage of 28 miles per gallon?

Solution

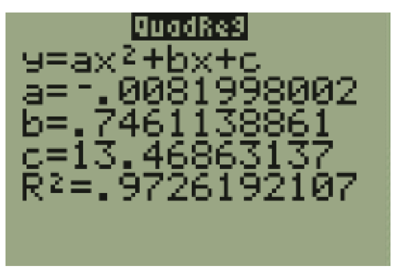
(a) Follow the steps outlined in the first example, and you should see the graph shown below.







(b) Under the `STAT` $\longrightarrow$ `CALC` menu, use the `QuadReg` option, leaving all the options unchanged:



The model found by the calculator is

$$\boxed{P_t = -0.008t^2 + 0.746t + 13.469}$$

(c) To make predictions for the mileage (P_t) based on speed (t), simply substitute those values into the model:

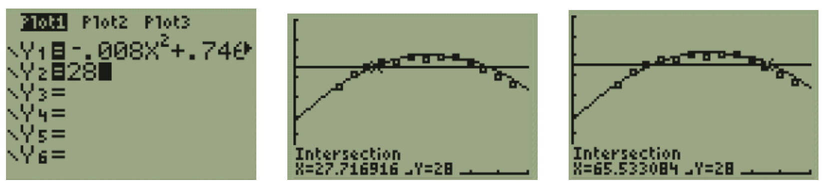
- At 62 mph: $P_t = -0.008(62)^2 + 0.746(62) + 13.469 = \boxed{28.97 \text{ mpg}}$
- At 90 mph: $P_t = -0.008(90)^2 + 0.746(90) + 13.469 = \boxed{15.81 \text{ mpg}}$

Which of these seems more likely to be reliable? The main difference between the two is that 62 mph falls within the range of speeds for which we have data, while 90 mph is faster than anything that was tested. Thus, we can call the first prediction (at 62 mph) an **interpolation** and the second an **extrapolation**. In general, interpolations are more reliable, because there may be effects we haven't observed outside the range we tested for.

(d) To find the speeds that match a mileage of 28 mpg, we need to solve

$$28 = -0.008t^2 + 0.746t + 13.469$$

Rather than doing this algebraically, we can use the same tool we used for this kind of problem in the last section: graph both sides of the equation and use the intersect tool on the calculator. Notice that there are two intersections this time, so we need to move the cursor close to the second intersection when we want to find that point.



The two speeds that we would expect to produce a gas mileage of 28 mpg are $\boxed{27.7 \text{ mph}}$ and $\boxed{65.5 \text{ mph}}$

TRY IT

Use the data and model from the previous TRY IT example.

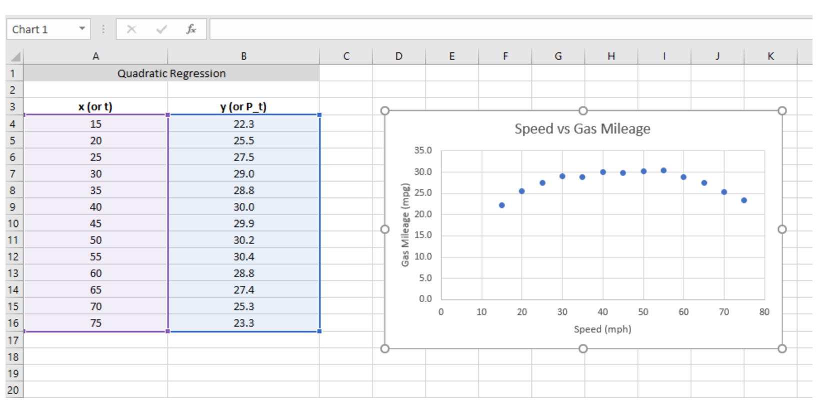
(a) Predict the values of y if $x = 3$ and if $x = 9$. Which prediction is more likely to be reliable?

(b) Find the value(s) of x that match $y = 10$ according to that model.

Using Excel

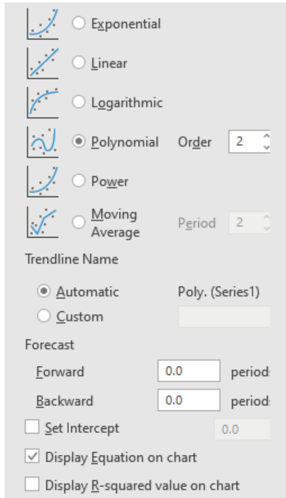
The process for using Excel with quadratic models is almost identical to the one shown in the last section for linear models. To illustrate, we'll use the dataset from the last example, with the relationship between speed and gas mileage.

First, record the data and insert a scatterplot, as before.

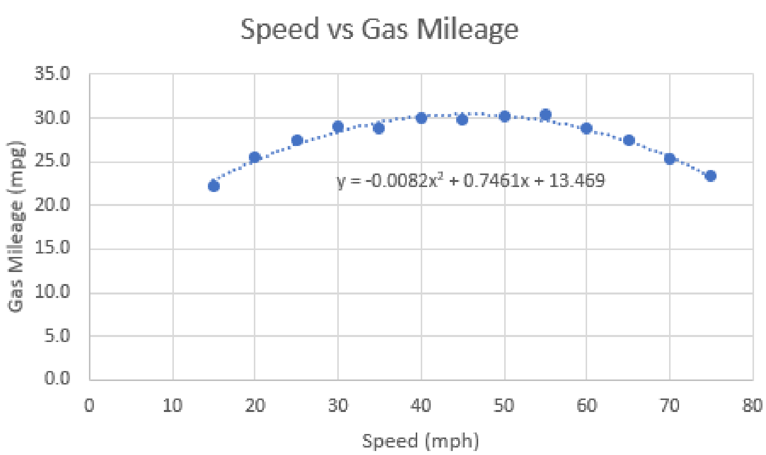

Quadratic Regression

x (or t)	y (or P_t)
15	22.3
20	25.5
25	27.5
30	29.0
35	28.8
40	30.0
45	29.9
50	30.2
55	30.4
60	28.8
65	27.4
70	25.3
75	23.3

Next, click the plus sign in the upper right corner of the chart, and at the bottom of the menu, click the arrow next to the "Trendline" option and select "More Options..."



This time, instead of selecting the linear option, select "Polynomial" with order 2 (that's the squared part of a quadratic model). As always, select the option to display the equation on the chart.



Exercises 1.2

1. The table below shows the distance that a baseball travels after being hit at various angles.

Angle (degrees)	Distance (feet)
10	115.6
15	157.2
20	189.2
24	220.8
30	253.8
34	269.2
40	284.8
45	285.0
48	277.4
50	269.2
58	244.2
60	231.4
64	180.4

(a) Use a graphing calculator or spreadsheet program to find a quadratic model that best fits this data, using angle as x and distance as y.

(b) Based on this model, what distance is expected for a ball hit at 55°?

(c) What distance is expected for a ball hit at 75°?

(d) Which of the two previous predictions is likely to be more reliable?

(e) What angle would you expect to yield a distance of 200 feet?

2. A ball is dropped from a height of a little over 5 feet, and the height is measured at small intervals. The table below shows the results.

Time (seconds)	Height (feet)
0.00	5.235
0.04	5.160
0.08	5.027
0.12	4.851
0.16	4.631
0.20	4.357
0.24	4.030
0.28	3.655
0.32	3.234
0.36	2.769
0.40	2.258
0.44	1.635

(a) Use a graphing calculator or spreadsheet program to find a quadratic model that best fits this data, using time as x and height as y.

(b) Based on this model, what height is expected after 0.30 seconds?

(c) What height is expected after 0.52 seconds?

(d) Which of the two previous predictions is likely to be more reliable?

(e) When do you expect the height of the ball to be 1 foot?

3. The table below shows the amount spent on movie theater tickets in the U.S. from 1997 to 2003.

Year	Spending (billions of dollars)
1997	6.3
1998	6.9
1999	7.9
2000	8.6
2001	9.0
2002	9.6
2003	9.9

(a) Use a graphing calculator or spreadsheet program to find a quadratic model that best fits the data. Let t represent the year, with $t = 0$ in 1997.

(b) Based on this model, how much would you expect to be spent on movie theater tickets in 2008?

(c) When would you expect movie theater ticket expenditure to reach \$11 billion?

4. The table below shows the number of FM radio stations in the U.S. from 1997 to 2003.

Year	Stations
1997	5542
1998	5662
1999	5766
2000	5892
2001	6051
2002	6161
2003	6207

(a) Use a graphing calculator or spreadsheet program to find a quadratic model that best fits the data. Let t represent the year, with $t = 0$ in 1997.

(b) Based on this model, how many FM stations would you expect there to be in 2010?

(c) When would you expect there to be 6500 stations?

5. The table below shows college textbook sales in the U.S. from 2000 to 2005.

Year	Textbook Sales (millions of dollars)
2000	4265
2001	4571
2002	4899
2003	5086
2004	5479
2005	5703

(a) Use a graphing calculator or spreadsheet program to find a quadratic model that best fits the data. Let t represent the year, with $t = 0$ in 2000.

(b) Based on this model, how much would you expect to be spent on college textbooks in 2015?

(c) When would you expect textbook sales to reach \$7 billion (\$7000 million)?

6. The table below shows the average amount of time spent per person on entertainment per year from 2000 to 2005.

Year	Hours
2000	3492
2001	3540
2002	3606
2003	3663
2004	3757
2005	3809

(a) Use a graphing calculator or spreadsheet program to find a quadratic model that best fits the data. Let t represent the year, with $t = 0$ in 2000.

(b) Based on this model, how many hours would you expect the average person to spend on entertainment in 2012?

(c) When would you expect the average amount of time to reach 4000?

SECTION 1.3 Exponential Models

In essence, linear models start with a simple assumption: growth happens based on adding a fixed amount over and over again. In this section, we'll investigate exponential models, which are the result of a different assumption, that growth happens by adding a *percentage* of the current total. In financial terms, this is the difference between simple interest (linear) and compound interest (exponential).

Many of the examples in this section deal with population growth, because this is a decent assumption to start with. Since the number of offspring in one generation depends on the number of parents in the previous generation, it makes sense that as the population grows, the number of offspring grows as well.

For instance, suppose that in order to predict the population of geese in a particular area, you assume that the number of geese will increase by 10% each year (accounting for births and deaths). According to this model, an initial population of 1000 geese would grow to 1100 geese at the end of one year, and the following year, 10% of 1100 would be added, making the growth speed up.

Notice that adding 10% to the total is the same as multiplying by 110%, or 1.1:

$$1000 + 1000(0.1) = 1000(1 + 0.1) = 1000(1.1)$$

We can track this for a few years, multiplying each year's population by 1.1 to get the next year's result, and watch a formula emerge.

Year 0: $P_0 = 1000$
Year 1: $P_1 = 1000(1.1)$
Year 2: $P_2 = 1000(1.1)(1.1)$
Year 3: $P_3 = 1000(1.1)(1.1)(1.1)$

It doesn't take long to recognize that each year's population will be 1000 multiplied by 1.1 over and over again. The number of times that 1.1 is multiplied is the same as the number of years that have elapsed since we started tracking the population.

We can write this more compactly by taking advantage of exponential notation, since for instance, $1000(1.1)(1.1)(1.1) = 1000(1.1)^3$.

In our example, then, we could predict the population of geese in any year using the following formula:

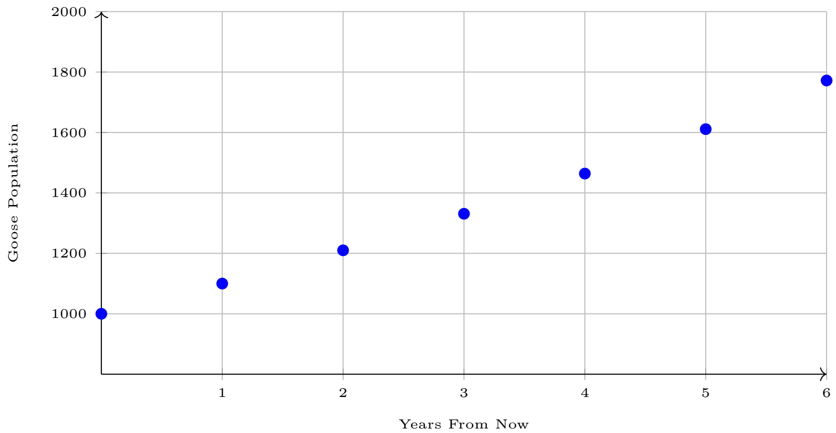
$$P_t = 1000(1.1)^t$$

The **growth rate** is 10%, and 1.10 is the **growth multiplier**. Each year's population is 1.10 times the previous year's population.

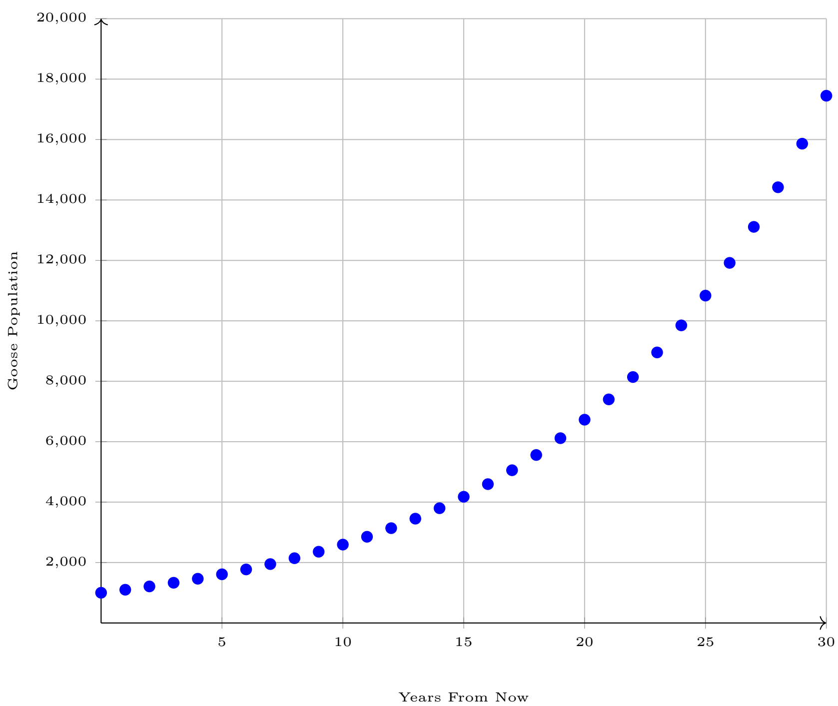
Year	Population	Growth from Previous Year
0	1000	
1	1100	100
2	1210	110
3	1331	121
4	1464	133
5	1611	147
6	1772	161

Notice that there is a constant *percentage* growth, so as the population increases, the number by which it grows gets larger each year.

If we plot these first few values, the graph is not quite linear, but it's not that far from a linear plot. Because of this, in the short term, linear models can approximate exponential models, even if it isn't a perfect fit.



As we begin to project further into the future, though, the model clearly deviates from a linear trend:



If the population had been growing linearly by 100 geese each year, the population at the end of 30 years would have only been 4,000 instead of nearly 18,000 under the exponential model. Most of this growth occurred in the second half; this is typical of exponential growth. Since the growth from one year to another depends on the size of the population, it grows much faster near the end, and the growth begins to snowball.

Exponential Growth

If a quantity starts at size P_0 and grows by $R\%$ (written as a decimal, r) every time period, then the quantity after t time periods is given by

$$P_t = P_0(1 + r)^t$$

The **growth rate** is r, and the **growth multiplier** is $1 + r$.

If r is negative, then instead of exponential growth there is **exponential decay**.

The growth multiplier is the common ratio between terms, and it can be used to recognize exponential growth from data, just like a common difference between terms can be used to recognize linear growth.

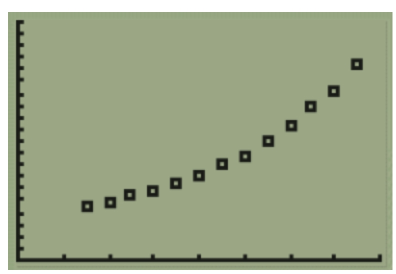
Year	Population	Ratio to Previous Year
0	1000	
1	1100	1.1
2	1210	1.1
3	1331	1.1
4	1464	1.1
5	1611	1.1
6	1772	1.1

Picking the Model Type How do we know whether to use a linear, quadratic, or exponential model (or some other type)? This is not a simple question, so in this chapter you will always be told what type of model to use.

However, we can look briefly at a simple example to see how this question may be answered. There are statistical tools that can be used to pick a model, but for now, we'll restrict ourselves to simply inspecting the data and doing a quick visual check.

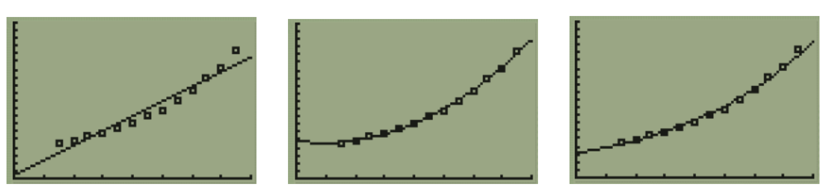
First, to distinguish between linear and exponential models, calculate the difference between each year's population and the next, and the ratio of the two populations. If the difference is roughly consistent, try a linear model. If, instead, the ratio is relatively unchanged, try an exponential model. Quadratic models are not as easy to observe this way, but in general, quadratic models fall somewhere between the other two.

Then, draw a scatterplot of the data and see if you can observe a trend. The more data you have available, the easier this is. Of course, this is not foolproof. For instance, consider the data graphed below.



This data doesn't necessarily look linear, but we could certainly find a linear model to approximate it, and it wouldn't be completely ridiculous.

We can compare the three types of models, and plot the results against the data:



Linear Quadratic Exponential

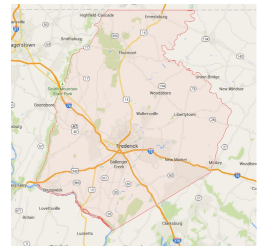
It looks like the quadratic and exponential models track the data better than the linear one, but between the two leaders, which is better? There's no clear answer, so we could probably pick either one and get decent results. In fact, trying out a bunch of models in the hopes of finding a perfect one can lead to *overfitting*, which happens when our entire focus is on explaining the data from the past, rather than being able to predict future results.

FREDERICK POPULATION

EXAMPLE 1

The population of Frederick County grew from 239,520 in 2012 to 241,409 in 2013, a growth of about 0.8%. If this growth rate continues, what is the population of Frederick County expected to be in 2025?

If $r = 0.008$, we can use the exponential growth formula to predict the population in 2025. To do so, however, we need to pick a year to be year 0. Since we're given the population in 2012 and 2013, we can use either one, but we'll choose 2013, so 2025 will be year 12.



$$\begin{aligned} P_{12} &= P_0(1+r)^t \\ &= 241,409(1+0.008)^{12} \\ &= \boxed{265,632} \end{aligned}$$

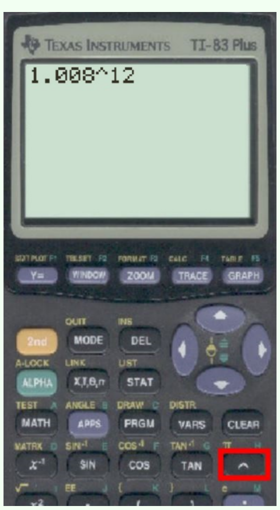
We expect the population of Frederick County to reach 265,632 by 2025.

TRY IT

India is the second most populous country in the world, with a population of about 1.252 billion in 2013. The population is growing by about 1.21% each year.

If this trend continues, what is India's population expected to grow to by 2030?

Using Your Calculator: Exponents



To evaluate expressions like 1.008^{12}, we'll use the exponent function on a calculator rather than multiplying 1.008 by itself 12 times. The exponent function is usually labeled like one of the following:

$$\boxed{\wedge} \quad \boxed{y^x} \quad \boxed{x^y}$$

To evaluate 1.008^{12}, we'd type 1.008 $\boxed{\wedge}$ 12 or 1.008 $\boxed{y^x}$ 12. Try it and make sure that you get an answer around 1.100338694.

EXAMPLE 2 TUITION PREDICTION

A friend is using the equation

$$P_t = 4600(1.072)^t$$

to predict the annual tuition at a local college. She says that the formula is based on years after 2010. What does this equation tell us?

Solution In this equation, $P_0 = 4600$, which is the initial tuition, so we infer that the tuition in 2010 is \$4600.

The growth multiplier is 1.072, so the growth rate is 0.072 or 7.2%. We expect tuition to grow by 7.2% each year.

In reality, we don't generally start a study already knowing the growth rate, as we did in the last few examples. Instead, we'll start with historical data, and we need to know how to use that to discover the growth rate.

We'll do this in a similar way to what we did in the section on Linear Models. If you remember, in that section, we had two ways to find the linear growth rate:

- Use two points from the data set (by default, the first and last points) to calculate the growth rate manually; this requires some algebra.
- Use all the points in the data set; use the calculator for this, since the details are complex.

With quadratic models, we defaulted to the second option and avoided the manual option altogether. However, we will see one example of finding r manually here, simply to see that it is doable. The algebra is a bit more complicated than it was for linear models, but still within grasp.

Finding the Growth Rate Manually

For a simple example, suppose that we knew that the population grew from 100 to 200 in 5 years. We can call 100 the initial population:

$$P_t = 100(1 + r)^t$$

The key is that we know a value of t and the corresponding population, P_t. If we substitute these into the equation, there will only be one unknown, r, and we can solve for it:

$$200 = 100(1 + r)^5$$

The steps to solve for r may be hard to follow at first, but you can use the exact same steps any time you need to solve one of these problems manually.

Remember that to solve an equation, we need to strip away everything except the piece we want. Here, we want to extract r, so we need to remove everything around it by undoing whatever operations apply. This also gives us an idea of the order, since we need to undo operations in the opposite order that we would simplify them according to PEMDAS (parentheses, exponents, multiplication, division, addition, subtraction).

Notice what's happening to r: it is added to 1 inside parentheses; since parentheses come first in the order of operations, we'll deal with this last as we solve for r. Next, this is raised to the fifth power, so we'll need to undo this second-to-last. Finally, the result is multiplied by 100, so we'll start by undoing that. Here's our process:

1. Undo multiplication by 100: divide both sides by 100
2. Undo raising to the 5th power: take the 5th root of both sides
3. Undo addition to 1: subtract 1 from both sides

That second step is the one that may be new to you, and it's likely the step that will give you the most trouble.

Let's see this in action.

CARBON DIOXIDE EMISSIONS

EXAMPLE 3

In 1990, the residential energy use in the US was responsible for 962 million metric tons of carbon dioxide emissions. By the year 2000, that number had risen to 1182 million metric tons. If the emissions grow exponentially and continue at the same rate, what will the emissions grow to by 2050?

The twist in this problem is that the growth rate is not explicitly given, so we'll have to find it before we can make our prediction.

We will let 1990 correspond to year 0, so 2000 is year 10.

Year	**Emissions (million tons)**
0	962
10	1182

We can put this information into the exponential growth model:

$$P_{10} = P_0(1+r)^{10}$$
$$1182 = 962(1+r)^{10}$$

Now we need to solve for r:

$$1182 = 962(1+r)^{10}$$

$$\frac{1182}{962} = (1+r)^{10} \qquad \text{Divide both sides by 962}$$

$$\sqrt[10]{\frac{1182}{962}} = 1+r \qquad \text{Take the 10th root of both sides}$$

$$\sqrt[10]{\frac{1182}{962}} - 1 = r \qquad \text{Subtract 1 from both sides}$$

$$r = \sqrt[10]{\frac{1182}{962}} - 1 = 0.0208 = 2.08\%$$

So if the emissions are growing exponentially, they are growing by about 2.08% per year. We can use this to predict the emissions in 2050, using 1990 as year 0:

$$P_{60} = 962(1+0.0208)^{60} = \boxed{2208.4 \text{ million metric tons of } CO_2 \text{ in } 2050}$$

TRY IT

The number of users on a social networking site was 45,000 in February when they officially went public, and grew to 60,000 by October. If the site is growing exponentially and growth continues at the same rate, how many users should they expect two years after they went public?

Rounding: If we had rounded the growth rate to 2.1%, our calculation for the emissions in 2050 would have been 3347. Rounding to 2% would have given a result of 3156. A very small difference in the growth rate gets magnified greatly in exponential growth. Thus, round the growth rate as little as possible, keeping at least three significant digits (numbers after any leading zeros). For instance, 0.41624 could be reasonably rounded to 0.416, and a growth rate of 0.001027 could be rounded to 0.00103.

Another note: Here we used two points to build an exponential model. If all we have is two data points, there's no reason necessarily to use an exponential model; a linear or quadratic model can fit two points just as well. This only makes sense when we have prior knowledge that convinces us that the growth in this instance will be exponential (like with population or things that are closely linked to it, such as pollution).

Using Your Calculator: Roots

In the previous example, we had to calculate the 10th root of a number. Many scientific calculators have a button for general roots that looks like:

$$\boxed{\sqrt[n]{\ }} \quad \boxed{\sqrt[x]{\ }} \quad \boxed{\sqrt[y]{x}}$$

To evaluate the 3rd root of 8, for example, we'd type either 3 $\boxed{\sqrt[n]{\ }}$ 8 or 8 $\boxed{\sqrt[n]{\ }}$ 3, depending on the calculator. Try it on yours to see—you should get 2.

If you can't find a general root button, you can use the property of exponents that

$$\sqrt[n]{a} = a^{1/n}.$$

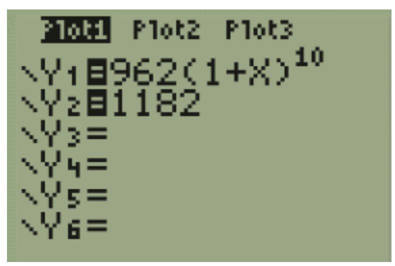
To compute $\sqrt[3]{8}$, then, you could use the exponent key on your calculator to evaluate $8^{1/3}$. Make sure that you use parentheses to preserve order of operations:

$$8 \ \boxed{y^x} \ (\ 1 \ \boxed{\div} \ 3\ )$$

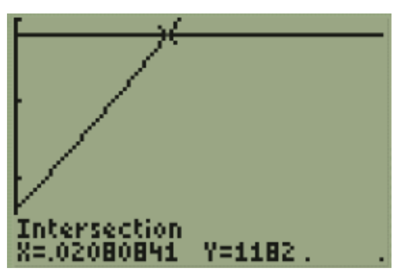
Finding the Growth Rate Using a Calculator

We can use the calculator for an example like the last one, without even resorting to regression (which we'll do a bit later). Remember that we used the intersect tool to solve for t in earlier models; we can do the same here to solve for r.

All we need is to graph both sides of the equation, using x instead of r:



Then use the intersect feature (press $\boxed{\text{2ND}}$ $\boxed{\text{TRACE}}$ to access the menu). Note that in order to visualize the intersection, we set the window for x values to be from 0 to 0.05 (since the growth rate is somewhere below 5% in this example) and y values between 900 and 1200. It may take some trial and error to get the window right, but this is not a crucial step of the process; the calculator can find an intersection even if it isn't visible.



Notice that the value the calculator gives for x is what we were looking for, since we used x to represent r: the growth rate is 2.08%, just as we found manually.

Now that you know how to do this, try going back to the first example, with the population of Frederick County, and try to calculate the growth rate that's given using either manual method or the calculator (or both, for practice).

Solving for Time

There's nothing new to say about solving for time using the calculator; we've already done this with the previous models in earlier sections, and we just used the same procedure to solve for the growth rate. However, we'll include an example here, just to remind you how this works.

SOLVING FOR TIME — EXAMPLE 4

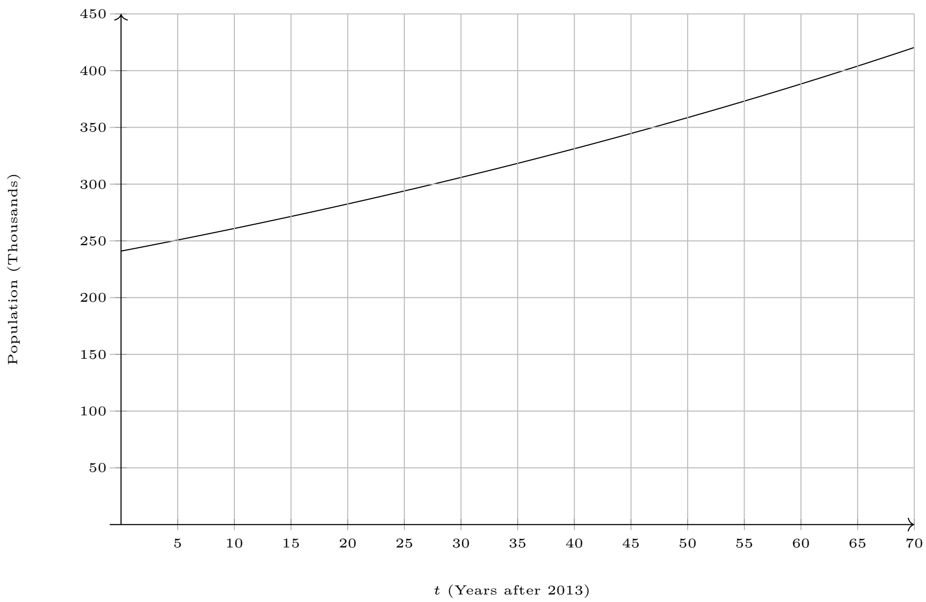
In the first example, we modeled the population growth of Frederick County from 2013 onward using the following equation:

$$P_t = 241,409(1 + 0.008)^t$$

Using this model, predict when the population will reach 400,000 people.

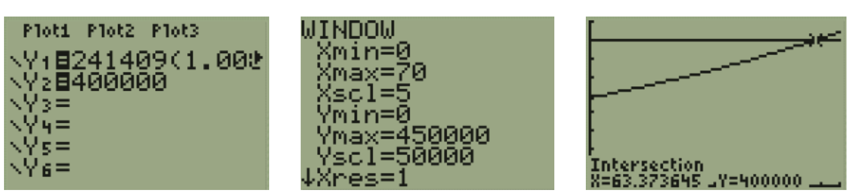
Solution

First, let's draw a graph to estimate the answer (an estimate like this may be all we need in some cases).



We can tell that the answer will be somewhere between 60 and 65 years after 2013. If we pick 63 as a reasonable guess, that means we expect the population to reach 400,000 around the year 2076.

To get a more precise guess (but remember, this is still an estimate, so there's no guarantee it will be more accurate than our first one), we can plot the model as well as the horizontal line at 400,000 and find their intersection:



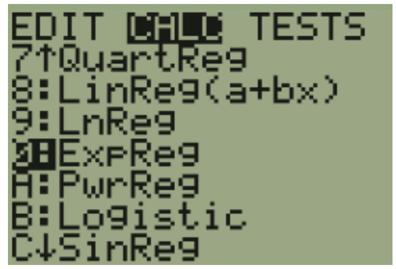
Notice that the intersection occurs around $x = 63.4$, so our initial guess of 63 was pretty close. Thus, we'll stick with 2076 as the year that we predict the population will reach 400,000.

TRY IT

Using the carbon emissions model from Example 3, predict when the emissions will reach 1500 million metric tons of carbon dioxide.

Exponential Regression

If we want to use more than two points to find an exponential model, we need to turn to the calculator. Thankfully, since we've already done linear and quadratic regression, there's not much to add here. Exponential regression can be found in the same menu, labeled `0: ExpReg`. Press the STAT button, then scroll to the right to the `CALC` menu:



The rest of the steps are familiar; start by entering data (with time in the first column and population in the second column), then access the `ExpReg` function, and leave the default options alone.

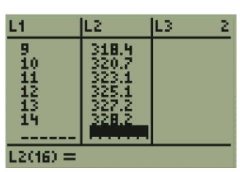
EXAMPLE 5 EXPONENTIAL REGRESSION

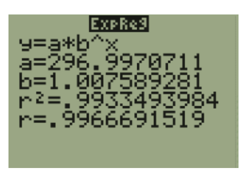
Build an exponential population model for the U.S. using data from 2005 to 2019.

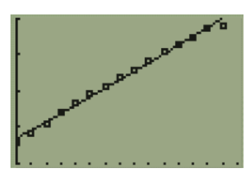
Solution We need to gather some data first; a quick Internet search yields the following population values:

Year	Population (in millions)
2005	295.5
2006	298.4
2007	301.2
2008	304.1
2009	306.8
2010	309.3
2011	311.6
2012	313.9
2013	316.1
2014	318.4
2015	320.7
2016	323.1
2017	325.1
2018	327.2
2019	328.2

Enter this into the calculator, using 2005 as year 0 (so 2019 will correspond to year 14), and use the exponential regression solver.







The exponential model is

$$\boxed{P_t = 297(1.0076)^t}$$

which means that the growth rate is 0.76%. We could then use this model to make predictions, as we've done in the last few examples.

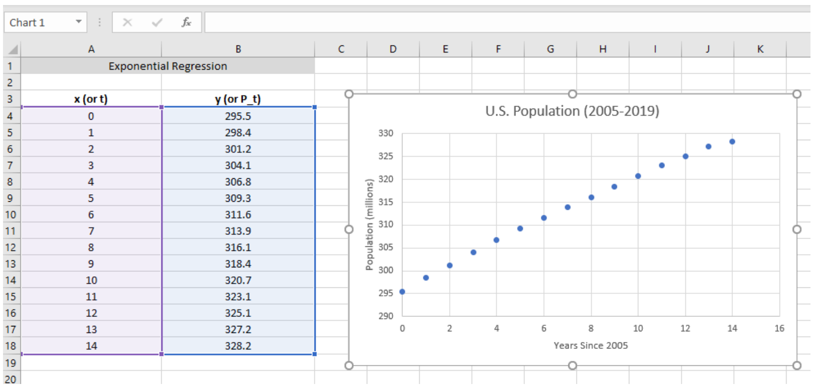
TRY IT

Pick another country, and build an exponential population model using data from the same years.

Using Excel

We've used Excel to build regression equations for linear and quadratic models; using the same process, you'll notice an option for an exponential model. We could do it that way; however, the form that Excel gives for the exponential model uses the natural base e, which we haven't addressed in this section. To avoid confusion, we'll use a different process to get a more familiar form; there's a built-in formula called `LOGEST`, which takes a range of y-values and a range of x-values, in that order, and returns an exponential model in the form $y = a(b)^x$, just like the graphing calculator.

First, record the data and insert a scatterplot, as before.

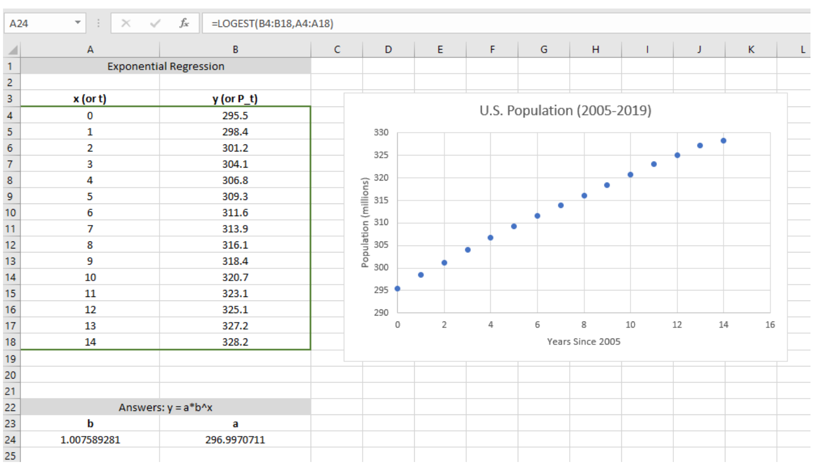


x (or t)	y (or P_t)
0	295.5
1	298.4
2	301.2
3	304.1
4	306.8
5	309.3
6	311.6
7	313.9
8	316.1
9	318.4
10	320.7
11	323.1
12	325.1
13	327.2
14	328.2

Next, select a cell (`A24` on the example spreadsheet) and enter the formula

```
=LOGEST(B4:B18,A4:A18)
```

This selects the values in the second column as y and the first column as x; the output is shown below.



x (or t)	y (or P_t)
0	295.5
1	298.4
2	301.2
3	304.1
4	306.8
5	309.3
6	311.6
7	313.9
8	316.1
9	318.4
10	320.7
11	323.1
12	325.1
13	327.2
14	328.2

b	a
1.007589281	296.9970711

Notice that the output of the formula places the answer for the base of the exponent in the selected cell, and the answer for the initial population in the next cell to the right.

The model given here is

$$P_t = 297(1.0076)^t$$

which is identical to the one from the graphing calculator.

Exercises 1.3

In problems 1–4, use a calculator to solve for the unknown variable, x or t.

1. $10^x = 5$

2. $7^t = 100$

3. $4(10^x) = 9$

4. $22(1.065)^{0.05t} = 37$

5. The population of the District of Columbia was approximately 572,000 in 2000, and has been growing at a rate of about 1.15%.

(a) Write an exponential model of the form $P_t = P_0(1+r)^t$ to describe the population of DC from 2000 onward.

(b) If this trend continues, what will DC's population be in 2025?

(c) When does this model predict that DC's population will reach 800,000?

6. Baltimore's population in 2010 was approximately 620 thousand, and has been decreasing at a rate of about 0.5% per year.

(a) Write an exponential model of the form $P_t = P_0(1+r)^t$ to describe the population of Baltimore from 2010 onward.

(b) If this trend continues, what will Baltimore's population be in 2030?

(c) When does this model predict that Baltimore's population will reach 500,000?

7. Diseases tend to spread exponentially. In the early days of AIDS, the growth rate was around 190%. In 1983, about 1700 people in the US died of AIDS. If the trend had continued unchecked, how many people would have died from AIDS in 2005?

8. The population of the world in 1987 was 5 billion and the annual growth rate was estimated at 2 percent. If the world population followed an exponential growth model, find the projected world population in 2015.

9. The population of Maryland was 5.17 million in 1999, and it grew to 6.05 million in 2019.

(a) Assuming that the population is growing exponentially, find the growth rate r for Maryland's population.

(b) Write an exponential model to describe the population of Maryland from 1999 onward.

(c) What is Maryland's population expected to be in 2030?

(d) When do you expect that Maryland's population will reach 8 million?

10. The population of Virginia was 6.87 million in 1999, and it grew to 8.54 million in 2019.

(a) Assuming that the population is growing exponentially, find the growth rate r for Virginia's population.

(b) Write an exponential model to describe the population of Virginia from 1999 onward.

(c) What is Virginia's population expected to be in 2022?

(d) When do you expect that Virginia's population will reach 10 million?

11. A bacteria culture is started with 300 bacteria. After 4 hours, the population has grown to 500 bacteria. If the population grows exponentially according to the formula $P_t = P_0(1+r)^t$,

(a) Find the growth rate r and write the full formula.

(b) If this trend continues, how many bacteria will there be in one day?

(c) How long will it take for this culture to triple in size?

12. A native wolf species has been reintroduced into a national forest. Originally 200 wolves were transplanted, and after 3 years, the population had grown to 270 wolves. If the population grows exponentially according to the formula $P_t = P_0(1+r)^t$,

(a) Find the growth rate r and write the full formula.

(b) If this trend continues, how many wolves will there be in ten years?

(c) If this trend continues, how long will it take the wolf population to double?

13. In 2009, the average compensation for CEOs in the U.S. was approximately \$10,800,000, and by 2016, this had risen to about \$12,800,000. By comparison, the average compensation for workers was \$54,700 in 2009 and \$55,800 in 2016. Assume that both values are growing according to an exponential model. Find the growth rate for both salaries; which is higher?

14. In 2008, approximately 131 million people voted in the U.S. general election, compared to about 139 million people in 2016. The total population of the U.S. was 304 million in 2008 and 323 million in 2016. Assume that both levels are growing exponentially. Find the growth rate for both populations; which is higher?

15. The table below shows the population of Canada from 2010 to 2019.

Year	**Population (millions)**
2010	34.0
2011	33.5
2012	34.7
2013	35.1
2014	35.4
2015	35.7
2016	35.1
2017	36.5
2018	37.1
2019	37.6

(a) Use a graphing calculator or spreadsheet program to build an exponential regression model, letting $t = 0$ in 2010.

(b) What does this model predict that the population of Canada will be in 2035?

(c) When does this model predict that Canada's population will reach 40 million?

16. The table below shows the population of Mexico from 2010 to 2019.

Year	**Population (millions)**
2010	114.1
2011	115.7
2012	117.3
2013	118.8
2014	120.4
2015	121.9
2016	123.3
2017	124.8
2018	126.2
2019	127.6

(a) Use a graphing calculator or spreadsheet program to build an exponential regression model, letting $t = 0$ in 2010.

(b) What does this model predict that the population of Mexico will be in 2040?

(c) When does this model predict that Mexico's population will reach 145 million?

SECTION 1.4 Logistic Models

There's an old legend about the inventor of chess, who, when he was asked by the ruler what reward he desired for his invention, told the ruler to simply pay him in rice. In fact, he said, the ruler could simply place one grain of rice on the first square of a chessboard, followed by two grains of rice on the next square, and so on, doubling each time. Surprised by how little the inventor requested, the ruler agreed, and began adding rice to the board.

He quickly ran into a problem, though. By the time he reached the 21st square, he had to deliver over a million grains of rice, and by the 31st square, it was over a billion squares. He called in his best mathematicians, and to his horror, they informed the ruler that if he kept up his side of the bargain, he would have to give the inventor a total of 18,446,774,073,709,551,615 grains of rice (about 1000 times the amount of rice harvested worldwide in recent years).

This exposes a flaw in using exponential models for population growth. Although doubling at every step corresponds to a 200% growth rate, even more modest rates exhibit the same behavior when observed over a long enough time: exponential models grow without bound, and they grow faster and faster over time.

Limited Resources

In 1950, the world population was 2.53 billion, and it grew to 5.32 billion by 1990, 40 years later. We can use these two data points to build an exponential model that predicts that t years from 1950, the world population in billions will be

$$P_t = 2.53(1.0188^t)$$

Let's test the model; let $t = 55$, for example, corresponding to the year 2005. The actual population in 2005 was 6.51 billion, and the model predicts that the population would be 7.03 billion. Not perfect, but not terrible either.

However, what about 2015? The model predicts 8.47 billion, and the actual population is only 7.32 billion. In other words, the model is getting worse, and it's consistently overestimating; in 2005 the estimate was only off by half a billion, but by 2015 the error has more than doubled. What's happening?

It turns out that the population growth rate is not constant, as the exponential model assumed. Instead, the growth rate is slowing. The world population grew from 1.60 billion in 1900 to 6.13 billion in 2000, so even if it grew by the same *amount* (not even the same *rate*, which would lead to even bigger numbers) we might naively assume that by 2100 the world population would top 11 billion. However, the United Nations estimates that the population will top 8 billion around 2050, and then *fall* back to around present-day levels by 2100.

What's going on here? There are many factors, but one of the most fundamental is that the earth has **limited resources**. Clearly, the population can't keep growing forever without bound; the earth cannot sustain a trillion people, for instance, given current technology, infrastructure, and access to food and water. This leads to our conclusion:

Exponential models are not good enough in the long term
because they don't account for limited resources.

Logistic Models

In the short term, exponential models can give decent estimates, but in the long run, they'll eventually give unsustainable results. To account for this, we turn to **logistic models**, which do account for limited resources.

Carrying Capacity

The **carrying capacity**, or **maximum sustainable population**, is the largest population that an environment can support.

Logistic Growth

If a population is growing in a constrained environment with carrying capacity M and growth rate r, then the population can be described by the logistic growth model:

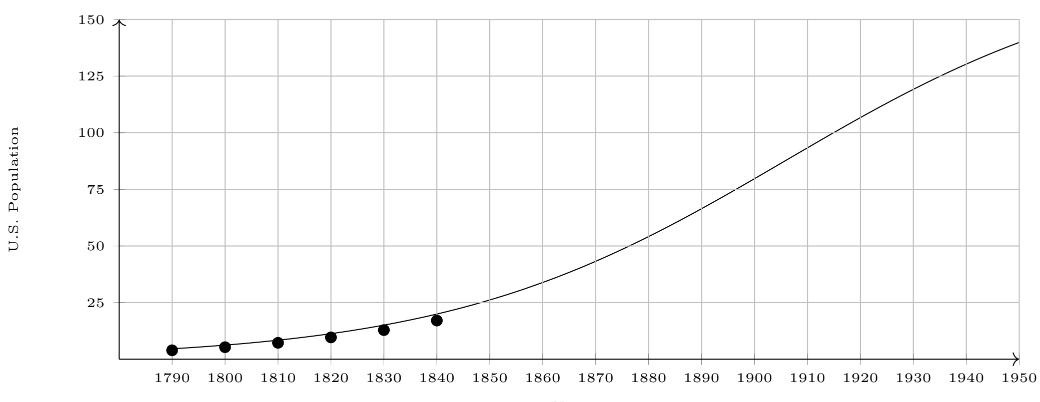
$$P_t = \frac{M}{1 + \left(\frac{M}{P_0} - 1\right) e^{-rt}}$$

This model of population growth is sometimes called the *Verhulst model*, after Pierre-François Verhulst, a Belgian mathematician who published the model in 1838[1] and used it in 1840 to predict the population of the U.S. up to 1940. His estimate of the 1940 population of the U.S. was off by less than 1%, a remarkable achievement.

He worked with the following data from 1790 to 1840:

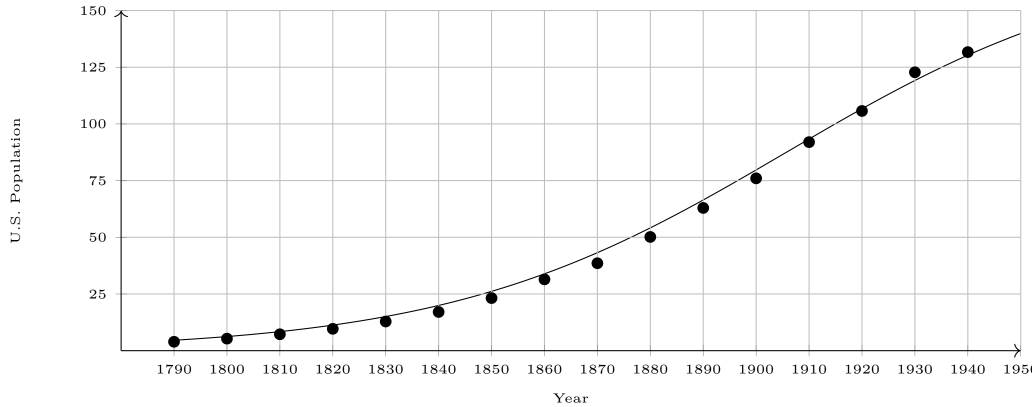
Date (Years AD)	**Population** (millions)
1790	3.929
1800	5.308
1810	7.240
1820	9.638
1830	12.866
1840	17.069

The graph below shows these data points, as well as the curve generated by a logistic model. Note the trademark S-shape of the curve; this is typical for logistic curves—they initially look like exponential curves, but then level off as the population approaches the carrying capacity.



The next graph shows the same model, but this time with data from the U.S. census filled in for the remaining years.

[1] It was rediscovered and re-derived several times over the following centuries by other mathematicians studying population growth.



Notice how closely the actual data tracks with the model's predictions; this is evidence of a good model, especially when we consider that the prediction was made ahead of time. More often than you might expect, would-be experts will reach for data from the past and magically find a model that fits the data nearly perfectly, but such models tend to fail at making predictions, and they don't actually offer any insight. A truly predictive model like this one with such accurate results is quite rare.

EXAMPLE 1 LIZARD POPULATION

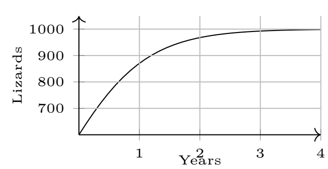
On an island that can support a population of 1000 lizards, there is currently a population of 600. These lizards have a lot of offspring and not many natural predators, so they have a very high growth rate of 150%. Use a logistic model to predict the lizard population 2 years from now.

Fill in the logistic model with the given information:

$$M = 1000,\ P_0 = 600,\ r = 1.5$$

$$P_t = \frac{M}{1 + \left(\frac{M}{P_0} - 1\right)e^{-rt}}$$

$$P_t = \frac{1000}{1 + \frac{2}{3}e^{-1.5t}}$$



Let $t = 2$ to predict the population in 2 years:

$$P_2 = \frac{1000}{1 + \frac{2}{3}e^{-1.5(2)}}$$

$$= \frac{1000}{1.0332} \approx \boxed{968 \text{ lizards}}$$

The model predicts that there will be approximately 968 lizards in 2 years.

TRY IT

A field contains 20 mint plants, and the number of plants increases at a rate of 70%, but the field can only support a maximum population of 300 plants. Use the logistic model to predict what the population will be in three years.

RABBIT POPULATION

EXAMPLE 2

A forest is currently home to a population of 200 rabbits. The forest is estimated to be able to sustain a population of 2000 rabbits, and the rabbits can grow at a rate of 50% per year.

(a) Find a model to predict the future rabbit population.

(b) Draw a graph of this model.

(c) Using this model, predict the population after 8 years.

(d) When will the population reach 1000 rabbits?

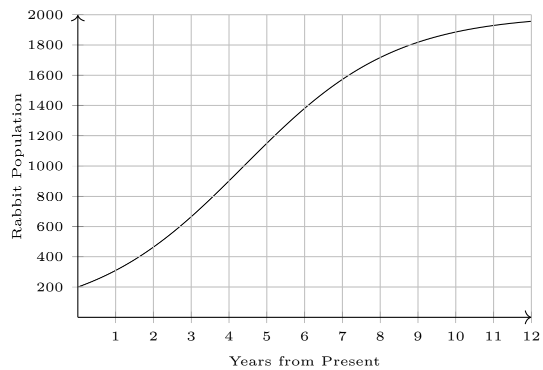
Solution

(a) We're told that $r = 0.5$, $M = 2000$, and $P_0 = 200$. Putting it all into the logistic model:

$$P_t = \frac{M}{1 + \left(\frac{M}{P_0} - 1\right) e^{-rt}}$$

$$\boxed{P_t = \frac{2000}{1 + 9e^{-0.5t}}}$$

(b) Graphing this equation:

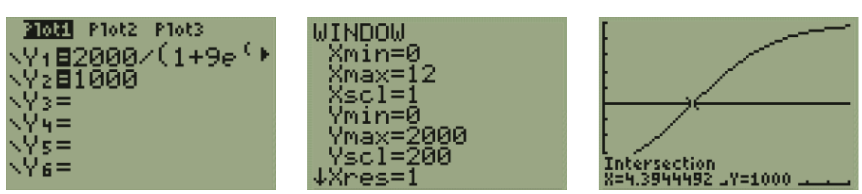


Note that according to the model, the rabbit population will level out near the carrying capacity in about 12 years.

(c) In 8 years, the population is predicted to be approximately

$$\begin{aligned} P_8 &= \frac{2000}{1 + 9e^{-0.5(8)}} \\ &= \frac{2000}{1 + 9(0.01832)} \\ &= \frac{2000}{1.1648} \qquad = \boxed{1717 \text{ rabbits}} \end{aligned}$$

(d) We'll use the calculator for this: graph the model and the horizontal line at 1000, and find their intersection:



The intersection is at $x = 4.394$, so the population is expected to reach 1000 rabbits after about $\boxed{4.4 \text{ years}}$

Logistic Regression

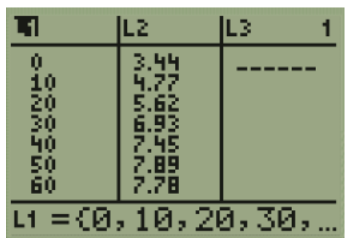
As before, it is unlikely that we will know at the beginning of a study what the growth rate is; it's more likely that we'll have data and need to build a model to fit. We'll use a graphing calculator to do this; although Excel can do it, it is not built-in to the menus that we have used before, and the process is very involved, so we will not include that in this chapter.

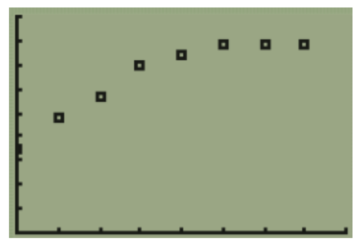
EXAMPLE 3 LOGISTIC REGRESSION

Build a logistic population model for New York City using the population data below.

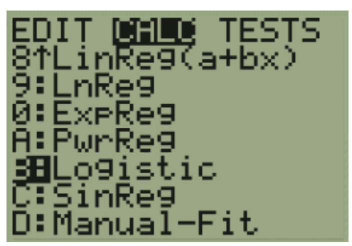
Year	Population (millions)
1900	3.44
1910	4.77
1920	5.62
1930	6.93
1940	7.45
1950	7.89
1960	7.78
1970	7.89

Solution First, enter the data (setting 1900 as year 0) and draw a scatterplot to get a visual sense of it:

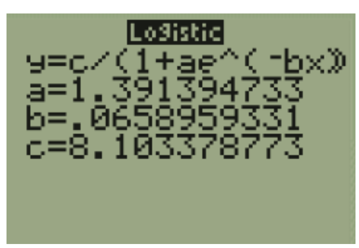


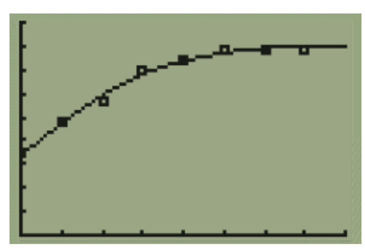


To get the regression equation, enter the $\boxed{\text{STAT}}$ `CALC` menu, then select `B: Logistic` near the bottom of the list. Once again, since we entered time in the first data column and population in the second, we can leave the default options unchanged.



The results are shown below; the graph on the right shows how the model tracks the data.





The model is

$$\boxed{P_t = \frac{8.1}{1 + 1.39e^{-0.066t}}}$$

Notice on the graph above how this model accounts for the leveling off of the population near the end.

Exercises 1.4

1. One hundred trout are seeded into a lake. Absent constraint, their population will grow by 70% a year. If the lake can sustain a maximum of 2000 trout, use a logistic growth model to estimate the number of trout after 2 years.

2. Ten blackberry plants started growing in a yard. Absent constraint, blackberries will spread by 200% a month. If the yard can only sustain 50 plants, use a logistic growth model to estimate the number of plants after 3 months.

3. A certain community consists of 1000 people, and one individual has a particularly contagious strain of influenza. Assuming the community has not had vaccination shots and are all susceptible, the spread of the disease in the community is modeled by

$$A = \frac{1000}{1 + 999e^{-0.3t}}$$

where A is the number of people who have contracted the flu after t days.

(a) How many people have contracted the flu after 10 days? Round your answer to the nearest whole number.

(b) What is the carrying capacity for this model? Does this make sense?

(c) How many days will it take for 750 people to contract the flu? Round your answer to the nearest whole number.

4. A herd of 20 white-tailed deer is introduced to a coastal island where there had been no deer before. Their population is predicted to increase according to

$$A = \frac{100}{1 + 4e^{-0.14t}}$$

where A is the number of deer expected in the herd after t years.

(a) How many deer will be present after 2 years? Round your answer to the nearest whole number.

(b) What is the carrying capacity for this model?

(c) How many years will it take for the herd to grow to 50 deer? Round your answer to the nearest whole number.

5. The table below shows the population of California from 2010 to 2019.

Year	Population (millions)
2010	37.3
2011	37.6
2012	38.0
2013	38.3
2014	38.6
2015	38.9
2016	39.2
2017	39.4
2018	39.5
2019	39.5

(a) Use a graphing calculator to build a logistic regression model, letting $t = 0$ in 2010.

(b) What does this model predict that the population of California will be in 2025?

(c) When does this model predict that California's population will reach 40 million?

(d) According to this model, what is the carrying capacity for California's population?

6. The table below shows the population of Florida from 2010 to 2019.

Year	Population (millions)
2010	18.7
2011	19.1
2012	19.3
2013	19.6
2014	19.9
2015	20.2
2016	20.6
2017	21.0
2018	21.2
2019	21.5

(a) Use a graphing calculator to build a logistic regression model, letting $t = 0$ in 2010.

(b) What does this model predict that the population of Florida will be in 2030?

(c) When does this model predict that Florida's population will reach 23 million?

(d) According to this model, what is the carrying capacity for Florida's population?