

## 1.2 Tables and graphs

Lung cancer, chronic bronchitis, bad breath, stains on your clothes, and the expense. These are just a few of the consequences of smoking cigarettes. With what we know now about the dangers of smoking, are people smoking more or less than they were ten years ago, fifty years ago, or even one hundred years ago?

Reality is, we don't have information on each individual person's smoking rate, so we can't answer this question exactly. We do have information on the total number of cigarettes sold each year. So maybe we should look at that total. Uh oh, that isn't going to work. There are way more people now than there were fifty or a hundred years ago. So, even if the same percentage of people smoke, and even if they each smoke the same amount as their predecessors did, we would have a much bigger number of cigarettes smoked now just because there are more people now.

Turns out a reasonable measure is to compare the number of cigarettes smoked per year *per person*. By taking into account the number of people we will be able to see whether people are smoking more or less, on average. That's what we want.

The table showing the smoking rate for select years. The smoking rate is the average cigarettes per year per person (adults).

Year	1900	1915	1930	1940	1950	1965	1975	1990	2000	2006
Smoking rate	54	285	1,485	1,976	3,552	4,258	4,122	2,834	2,049	1,619

Source: Center for Disease Control and Prevention

To make sense of these numbers, suppose there are five friends. Three don't smoke at all, so that is 0 cigarettes in a year. Another smokes only occasionally, maybe 100 cigarettes a year. The fifth smokes "a pack a day," which adds up to 7,300 cigarettes in a year because

$$\frac{1 \text{ pack}}{\text{day}} * \frac{20 \text{ cigarettes}}{\text{pack}} * \frac{365 \text{ days}}{\text{year}} = 20 \times 365 = \frac{7,300 \text{ cigarettes}}{\text{year}}$$

(Not sure about this calculation? Not to worry. More about unit conversions in §1.4.) These five people smoke a total of

$$0 + 0 + 0 + 100 + 7,300 = 7,400 \text{ cigarettes per year}$$

so when we divide by the number of people we get

$$\frac{7,400 \text{ cigarettes per year}}{5 \text{ people}} = 7,400 \div 5 = 1,480 \text{ cigarettes per year per person}$$

which is less than the average of 1,619 cigarettes per year per person for 2006 (the last year the CDC published the data), which is the way things seem to be heading.

We can tell a lot of information from this table. For example, what was the smoking rate in 1965, and how does that compare to 2006? The answers appears in the table, a whopping 4,258 cigarettes per person in 1965 and 1,619 cigarettes per person in 2006.

When did the consumption first pass 3,000? That answer does not appear in the table, but we can use the information in the table to make a good guess. In 1940, there were an average of 1,976 cigarettes per person per year and by 1950, there were 3,552. Somewhere between 1940 and 1950 the number first climbed above 3,000. More specifically, the number we're looking for (3,000) is a lot closer to the 1950 figure (3,552) than to the 1940 figure (1,976). So, it would be reasonable to guess close to 1950. I'd say 1947. Of course, you might guess 1946 or 1948, or even 1949 and those would be good guesses too.

When did the consumption drop below 3,000 again? This answer also does not appear in the table, but falls somewhere between 1975 when consumption was 4,122 and 1990 when consumption was 2,834. Here I'd guess just before 1990, say in 1989.

What's changing are the number of cigarettes smoked per person per year and the year. Those are our variables. The smoking rate is a function of year, and it's what we care about, so it's the dependent variable. Time, as measured in years, is the independent variable.

$S$  = smoking rate (cigarettes per year per person)  $\sim$  dep

$Y$  = year (years since 1900)  $\sim$  indep

Quick note on how we deal with actual years. Since the year 0 doesn't make sense in this problem, it is convenient to measure time in years since 1900, since that's the earliest year mentioned. Officially we should rewrite our table as:

$Y$	0	15	30	40	50	65	75	90	100	106
$S$	54	285	1,485	1,976	3,552	4,258	4,122	2,834	2,049	1,619

Notice where the variable names are listed in the table. In a horizontal format like this table, the independent variable ( $Y$ ) is in the top row, with the dependent variable ( $S$ ) is in the bottom row. If you want to write your table in a vertical format, that's okay too. Just put the independent variable in the left column, with the dependent variable in the right column. It might help to remember that the independent variable goes first (either top or left) and the dependent variable follows (either bottom or right).

Where the variables go in a table is not something you can figure out. It's a **convention** – a custom, practice, or standard used within the mathematical community. Though based on reason, it often involves some arbitrary choice, which is why we can't figure it out. So, whenever some practice is introduced to you as a “convention,” you need to memorize it.

Horizontal table format:

indep				
dep				

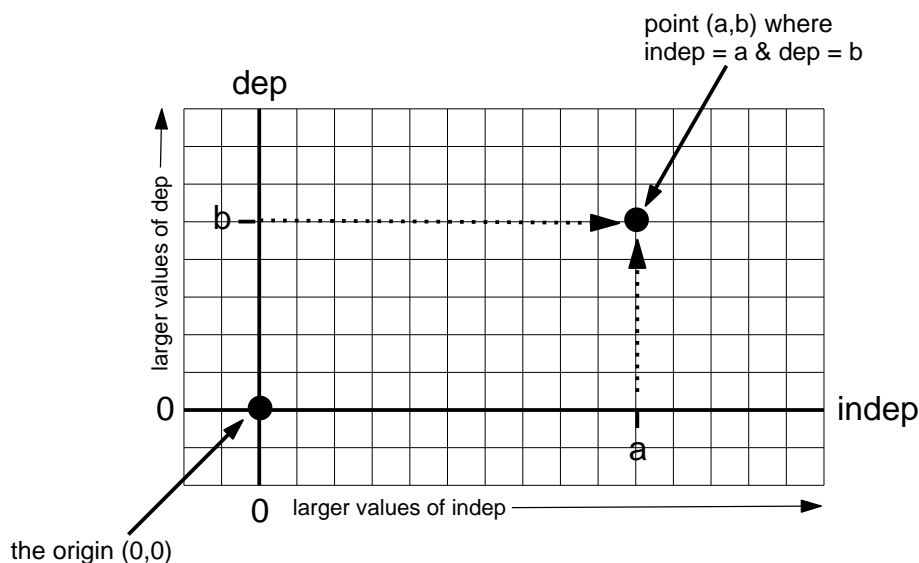
Vertical table format:

indep	dep

Tables are useful because they contain specific numbers, but it can be difficult to guess or see general trends. For that, a picture is worth a thousand words – or numbers, in this case. By “picture” we mean a graph of the function.

Throughout this text, we draw graphs by hand. On graph paper. Seriously. You might wonder why we do that when graphing calculators, spreadsheet programs, graphing “apps,” or computer algebra systems all can draw graphs for us. The answer is we want to understand graphs better, and I promise that drawing them by hand will help you do that. (Different instructors have different opinions on the importance of graphing by hand, so be sure to ask your instructor what you are expected to do. Even if you’re allowed to use some type of graphing technology, I strongly encourage you to practice drawing graphs by hand as well.)

There is a standard set-up for a graph.



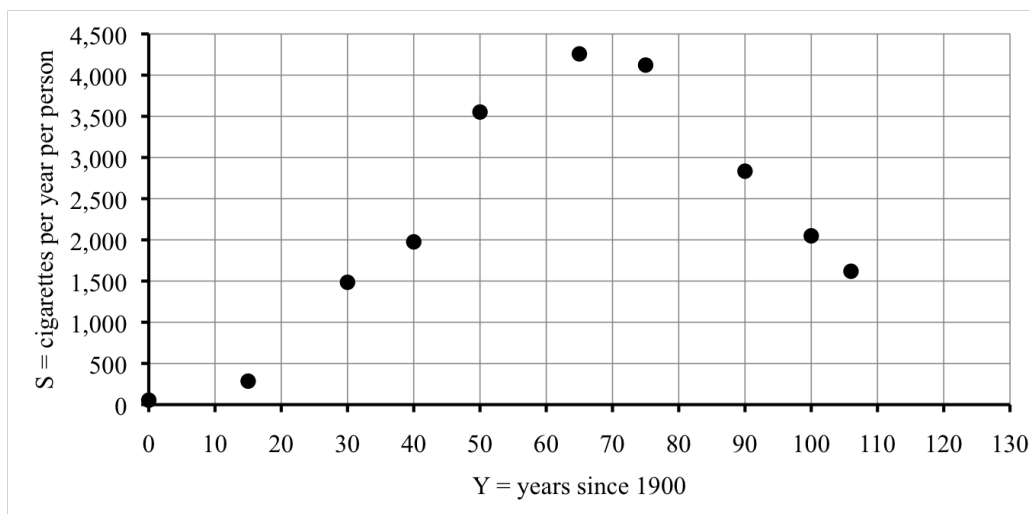
The graph is based on a horizontal line and vertical line, called the **axes**. Where they cross is a point called the **origin**. It represents where each variable is 0. By convention, the independent variable is measured along the horizontal axis, with larger values progressing to the right of the origin, and negatives to the left. Similarly, the dependent variable is

measured along the vertical axis, with larger values progressing up from the origin, and negatives down. Each gridline counts the same number, called the **scale**, but the scale for the vertical may be different from the scale for the horizontal. Each pair of values of the independent and dependent variable from our table correspond to a point on our graph.

In the graph of smoking rates, the independent variable is  $Y$ , the year, so that goes on the horizontal axis for our graph. Our dependent variable is  $S$ , the smoking rate, so that goes on the vertical axis. For the scale, it works nicely to count by 10s for years and count by 500s for the smoking rate.

There's a certain amount of guess and check involved in figuring out a good scale for each axis. As a general rule of thumb we would like the graph to be as large as possible so we can see all of its features clearly. But, not so big that it runs off the graph paper. What matters is that the gridlines are evenly scaled and that they can handle large enough numbers. Speaking of which, it's a good idea to leave a little room to extend the graph a little further than the information we have in the table, in case we get curious about values beyond what we have already.

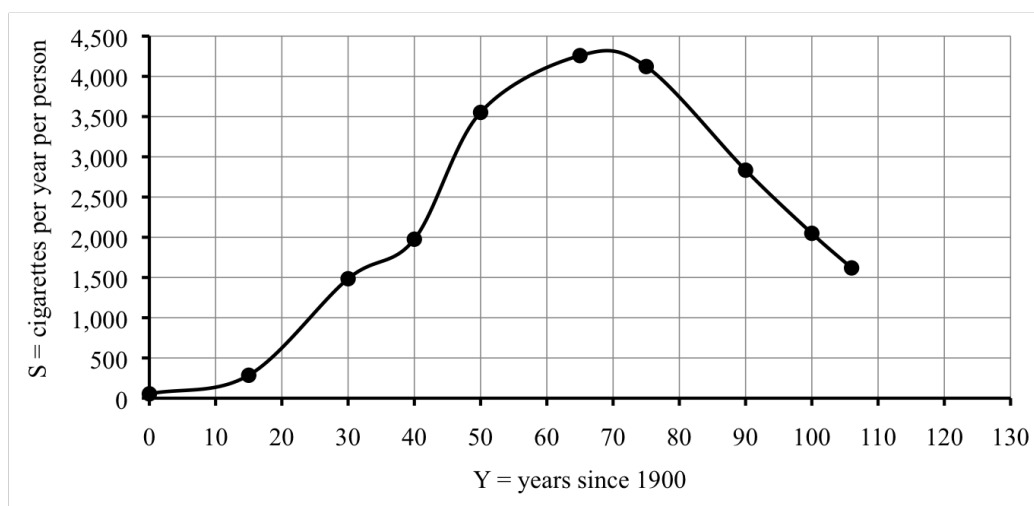
With realistic numbers it's normal to have numbers in the table that are not exactly where the gridlines are. It is very helpful to count by round numbers (2s, 5s, 10s, etc.) because it makes guessing in between easier. Easier for you drawing the graph. Easier for someone reading your graph.



To plot each point, we start at the origin and move right to that  $Y$ -value, and then up to that  $S$ -value. When a value doesn't land exactly on a grid mark, we have to guess in between. For example, in 1900, when  $Y = 0$  so we don't move right at all, just up to  $S = 54$ . The first labeled gridline on our graph is 500. Where's 54? It's between 0 and 500, very close to 0. Our point is just a tiny bit above the origin. In 1915 we have  $Y = 15$ .

Our labeled gridlines are for 10 and 20, so 15 must land halfway in between. The smoking rate to 285, which is around halfway between 0 and 500. Etc.

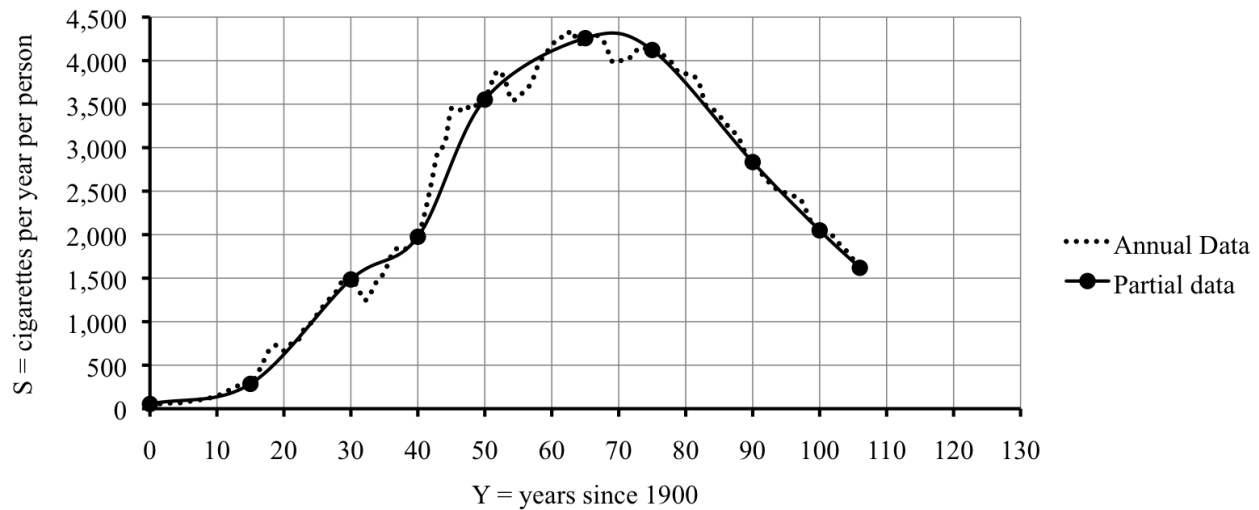
What we have so far is a **scatter plot** of points. Can you see why it's called that? Anyway, our whole goal here was to be able to understand smoking rates better by having a graph. You may already begin to see a curve suggested by the points. Time to draw it in. I don't mean drawing a line between each pair of points, like you do in the children's game "connect the dots." That isn't quite right. It was probably more of a continuous trend and so the graph should be smoother.



When we draw in this smooth curve for the graph, what we are really doing is making a whole lot of guesses all at once. For example, from the table we guessed that the smoking rate passed 3,000 in around 1947, and dropped back to that level in around 1989. What does the graph show? If we look where the horizontal gridline for 3,000 crosses our graph, it crosses in two places. First, between the vertical gridlines for 40 and 50, and perhaps slightly closer to 50. I'd say  $Y \approx 47$ , in the year 1947. Sure. The second time is between the gridlines for 80 and 90, much closer to 90. Looks like  $Y \approx 88$ , in the year 1988. We guessed 1989. Close enough.

Don't forget that when we drew in that curve it was really just a guess. We're sure about the points we plotted, but we're only guessing about where to draw the curve in. That means we're not sure about the other points. If we knew a lot more points we could have a more accurate graph.

Turns out more data is available from the CDC. The full table of data from the CDC shows that consumption first topped 3,000 as early as 1944. Here's an example where the history tells you more than the mathematics as cigarette consumption rose sharply during World War II. Our guess about 1988 or 1989 was spot on. Look at how the graph from the full data (the dotted line) compares to our guess.



## Homework

Start by doing Practice exercises #1-4 in the workbook.

Do you know ...

- Where the independent and dependent variables appear in a table and in a graph?
- How to guess values from a table or from a graph?
- How to make a graph from a table?
- Why we start each axis at 0?
- What we mean by scaling an axis evenly?
- How to make a table and then a graph from a story?
- Why we draw in a smooth line or curve connecting the points?
- What type of graphing technology, if any, you're allowed to use? *Ask your instructor.*

**If you're not sure, work the rest of exercises and then return to these questions. Or, ask your instructor or a classmate for help.**

## Exercises

5. The table lists estimates of Earth's population, in billions, for select years since 1800.

Year	1800	1850	1900	1950	1970	1990	2000
Population	.98	1.26	1.65	2.52	3.70	5.27	6.06

Source: "The World at Six Billion" United Nations report

*Story also appears in 1.3 Exercises*

Use the table to find or reasonably guess the answers to the following questions.

- What was the population of Earth in 1850?
  - What do you think the population of Earth was in 1860?
  - What do you think the population of Earth was in 1960?
  - In what year do you think the population of Earth first exceeded 2 billion?
  - In what year do you think the population of the world will exceed 7 billion?
  - Identify the variables, including units and dependence.
6. Your local truck rental agency lists what it costs to rent a truck (for one day) based on the number of miles you drive the truck.

Distance driven (miles)	50	100	150	200
Rental cost (\$)	37.50	55.00	72.50	90.00

*Story also appears in 1.3 and 4.4 Exercises*

Use the table to find or reasonably guess the answers to the following questions.

- How much does it cost to rent a truck if you drive it 100 miles?
- How many miles did you drive a truck costing \$90.00 to rent?
- If you rent a truck and drive it 75 miles, how much do you think it will cost?
- If you rent a truck and drive it 10 miles, how much do you think it will cost?
- If you rent a truck and it costs \$60.00, about how many miles was it driven?
- Identify the variables, including units, realistic domain, and dependence.
- Draw a detailed graph illustrating the dependence based on the points given in the table. Be sure your axes are labeled and evenly scaled. Sketch in a smooth curve connecting the points.
- Use your graph to check your answers to the questions. Modify, if necessary.

7. The temperature was  $40^{\circ}\text{F}$  at noon yesterday downtown Minneapolis but it dropped  $3^{\circ}\text{F}$  an hour in the afternoon. *Story also appears in 1.1 and 4.1 Exercises*
- Make a table of reasonable values.
  - Draw a graph illustrating the dependence. Count time in hours after noon.
  - According to your table and graph, when did the temperature drop below freezing ( $32^{\circ}\text{F}$ )?
  - According to your graph, when did the temperature drop below  $0^{\circ}\text{F}$ . Does that seem realistic? *Here in the midwest there are no oceans or mountains to moderate large temperature changes.*
8. Mrs. Nystrom's Social Security benefit was  $\$746.17/\text{month}$  when she retired from teaching in 2009. She had taught in elementary school since I was a girl. Benefits have increased by 4% per year. *Story also appears in 1.1 and 5.1 Exercises*
- Make a table of reasonable values using  $N$  for Mrs. Nystrom's benefits (in dollars) and  $Y$  for time (in years since 2009).
  - Draw a graph illustrating the dependence. Scale your graph to show up through the year 2020 and  $\$1,200$ .
  - According to your table and graph, when did her benefit pass  $\$900/\text{month}$ ?
  - If you extend your graph to 2020, what would you estimate Mrs. Nystrom's benefit will be then, assuming these increases continue?
9. The table adapted from shows the "heat index" as a function of humidity at an air temperature of  $88^{\circ}\text{F}$ . With up to about 40% humidity,  $88^{\circ}\text{F}$  feels like it's  $88^{\circ}\text{F}$ . But if the humidity rises to 60%, then it feels like it is  $95^{\circ}\text{F}$ ; that is, the heat index is  $95^{\circ}\text{F}$ .

Humidity (%)	50	60	70	85	90	95
Heat index ( $^{\circ}\text{F}$ )	91	95	100	110	113	117

Source: National Oceanic and Atmospheric Administration

All of the following questions refer to situations when the air temperature is  $88^{\circ}\text{F}$ .

- What is the heat index when the humidity level is 70%?
- At what humidity level does it feel more like  $98^{\circ}\text{F}$ ?
- Heat exhaustion is likely to occur when the heat index reaches  $105^{\circ}\text{F}$ . At what humidity level will heat exhaustion likely occur?
- The heat index is considered danger in the range from  $105^{\circ}\text{F}$  to  $129^{\circ}\text{F}$ . What range of humidity levels are considered dangerous?



- (e) What do you think the heat index would be at 99% humidity?
- (f) Identify the variables, including units, realistic domain, and dependence.
- (g) Draw a detailed graph illustrating the dependence based on the points given in the table. Be sure your axes are labeled and evenly scaled. Sketch in a smooth curve connecting the points.
- (h) Use your graph to check your answers to the questions. Modify, if necessary.