3 DIFFERING ROLES OF TABLES AND GRAPHS

Tables and graphs are the two fundamental vehicles for presenting quantitative information. They have developed over time to the point that we now thoroughly understand which works best for different circumstances and why. This chapter introduces tables and graphs and gives simple guidelines for selecting which to use for your particular purpose.

Tables and graphs are the two primary means to structure and communicate quantitative information. Both have been around for quite some time and have been researched extensively to hone their use to a fine edge of effectiveness. The best practices that have emerged are easy to learn, understand, and put to use in your everyday work with numbers.

Occasionally, the best way to display quantitative information is not in the form of a table or graph. When the quantitative information you want to convey consists only of a single number or two, written language is an effective means of communication; your message can be expressed simply as a sentence or highlighted as a bullet point. If your message is that last quarter's sales totaled $1,485,393 and exceeded the forecast by 16%, then it isn't necessary to structure the message as a table, and there is certainly no need to create a graph. You can simply say something like:

Q2 sales = $1,485,393, exceeding forecast by 16%

Alternatively, it wouldn't hurt to structure this message in simple tabular form such as this:

Q2 Sales	Compared to Forecast
$1,485,393	+16%

FIGURE 3.1 This table shows sales value information, arranged in columns.

or this:

Q2 Sales	$1,485,393
Compared to Forecast	+16%

FIGURE 3.2 This table shows sales value information, arranged in rows.

If you're like a lot of folks, however, you might be tempted while structuring this information as a table to jazz it up in a way that actually distracts from your simple, clear message—perhaps something like this:

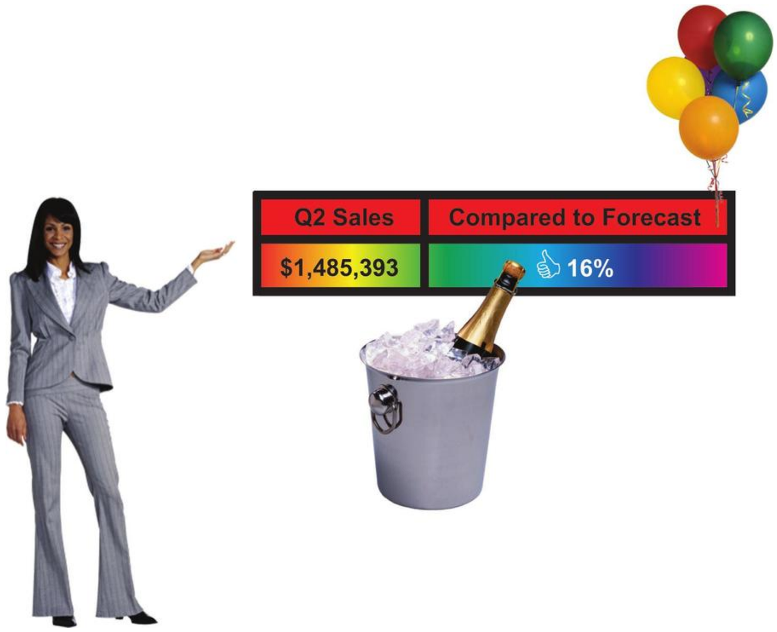

FIGURE 3.3 This table shows sales values, designed to impress, or perhaps to entertain, but not primarily to communicate.

You might even be tempted to pad the report with an inch-thick stack of pages containing the details of every sales order received during the quarter, eager to demonstrate how hard you worked to produce those two sales numbers. However, as we'll observe many times in this book, there is *eloquence in simplicity*.

Quantities and Categories

Before we launch into an individual examination of tables and graphs, let's review an important fact, common to both, that quantitative messages are made up of two types of data:

- Quantitative
- Categorical

Quantitative values measure something (number of orders, amount of profit, rating of customer satisfaction, etc.). Categorical items (i.e., members of a category) identify what the quantitative values measure. These two types of data fulfill different roles in tables and graphs. In the following simple table, which displays exempt and non-exempt employee compensation by department, all the information consists either of quantities being measured (compensation) or items belonging to categories (sales, operations, and manufacturing are items in the department category) to which the quantities relate.

Department	Exempt	Non-Exempt
Sales	950,003	1,309,846
Operations	648,763	2,039,927
Manufacturing	568,543	2,367,303
Total	$2,167,309	$5,717,076

FIGURE 3.4 This is a table of employee compensation information that you can use to practice distinguishing quantitative and categorical data.

The labels that identify the various departments in the far left column, including Total, belong to the category Department, and the labels Exempt and Non-exempt belong to a separate category that we could call Employee Type. All of the other data (i.e., all the numbers in this table) are quantitative values.

Do numbers always represent quantitative values? No. Sometimes numbers are used to label things and have no quantitative meaning. Order numbers (e.g., 1003789), numbers that identify the year (e.g., 2011), and numbers that sequence information (item number 1, item number 2, item number 3, etc.), are examples of numbers that express categorical data. One simple test to determine this distinction is to ask the question, “Would it make sense to add these numbers up, or to perform any mathematical operation on them?” For instance, would it make sense to add up order numbers? No, it wouldn’t. How about numbers that rank a list of sales people as number 1, number 2, etc.? Once again, the answer is no; therefore, these numbers represent categorical, rather than quantitative, data.

Let’s take a look at one more example to practice making the distinction between categorical and quantitative information. Using the graph below, test your skills by identifying which components represent categorical items and which represent quantitative values.

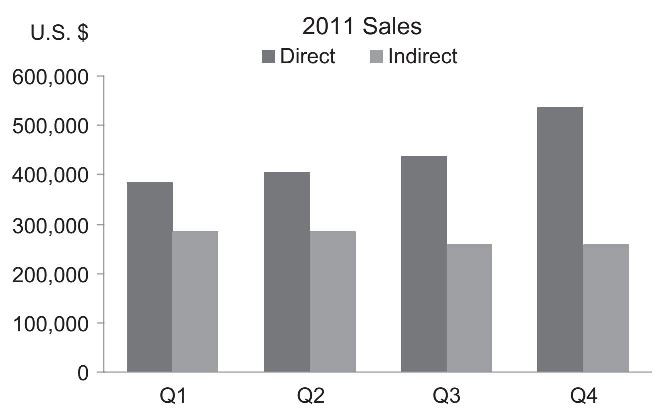

FIGURE 3.5 This graph depicts sales information for practice in distinguishing quantitative and categorical data.

Let’s examine one component at a time. For each of the items below, indicate whether the information is categorical or quantitative.

1. The values of time along the bottom (Q1, Q2, etc.)
2. The dollars along the left side
3. The legend, which encodes Direct and Indirect
4. The vertical bars in the body of the graph
5. The title in the top center

Once you’ve identified each, take a moment to compare your answers to those on the right.

ANSWERS
1. Categorical, labeling the quarters of the year
2. Quantitative, providing dollar values for interpreting the heights of the bars
3. Categorical, providing a distinction between direct and indirect sales
4. Both quantitative and categorical; the heights of the bars encode quantitative information about sales in dollars; the colors of the bars encode categorical data identifying which sales are direct vs. indirect
5. Categorical, identifying the year of the sales

• • • • • • •

Did you catch the dual role of the bars, that they contain both quantitative values and categorical items? With a little practice, you will be able to easily deconstruct graphs into quantitative and categorical data. This ability will enable you to apply the differing design practices that pertain to each type of data.

Choosing the Best Medium of Communication

Choosing whether to display data in one or more tables, one or more graphs, or some combination of the two, is a fundamental challenge of data presentation. This decision should never be arbitrary. It is sad, however, that this choice is often made using the "eeney, meeney, miney, moe" method. Imagine that you're Joe, and you've just interviewed three candidates for a new position in your department. It's now time to report the results to your boss, who's responsible for making the hiring decision. During interviews with the candidates, you used your company's handy interview score sheet to evaluate each person's aptitudes on a zero to five point scale in six areas: experience, communication, friendliness, subject matter knowledge, presentation, and education. How do you display the results? Well, if you're like a lot of people who feel pressure to make an impression, you might decide to use something unusual, like one of those radar charts that are available in your spreadsheet software. As you hand your boss the following report, you hope that he's thinking: "Wow, that Joe is an exceptional employee. Perhaps he deserves a raise."

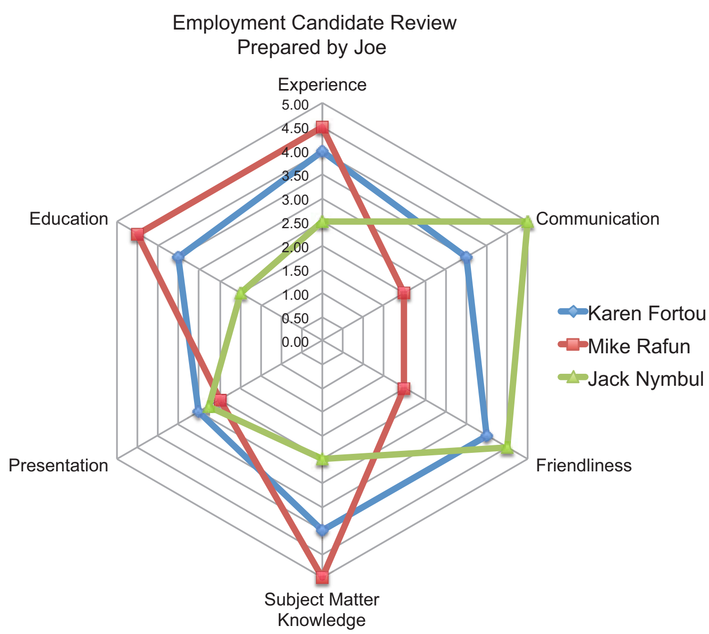

FIGURE 3.6 This radar chart presents aptitudes of job candidates in an overly complicated way.

What's sad is that the boss, upon first glance, might actually think, "That Joe certainly outdid himself," despite the fact that he hasn't a clue how to read this spider-web of a chart. He might actually blame himself for lacking the skill that's needed to read this chart, assuming he must have missed that day in math class when they covered this. Regardless of how the boss responds, Joe would have made his job easier if he'd prepared something like this instead:

Employment Candidate Review

	Candidates		
Rating Areas	Karen Fortou	Mike Rafun	Jack Nymbul
Experience	4.00	4.50	2.50
Communication	3.50	2.00	5.00
Friendliness	4.00	2.00	4.50
Subject matter knowledge	4.00	5.00	2.50
Presentation	3.00	1.50	2.75
Education	3.50	4.50	2.00
Average Rating	3.67	3.25	3.21

FIGURE 3.7 This table presents the same information as Figure 3.6 but in a way that is simple and clear.

This table presents the information in a way that is simple to understand and efficient to use. To select the appropriate medium of communication, we must understand the needs of our audience as well as the purposes for which various forms of display can be effectively used. Let's begin by considering the overall strengths and weaknesses of tables and graphs.

Tables Defined

A table is a structure for organizing and displaying information; a table exhibits the following characteristics:

- Information is arranged in columns and rows.
- Information is encoded as text (including words and numbers).

A single set of values, occupying a single column or row, is merely a list, not a table.

Although the column and row structure of tables is often visually reinforced by grid lines (i.e., horizontal and vertical lines outlining the columns and rows), it is the arrangement of the information that characterizes tables, not the presence of lines that visibly delineate the structure of the underlying grid. In fact, as we will see later in the chapter on table design, grid lines must be used with care to keep them from diminishing a table's usefulness.

Tables are not used exclusively to display quantitative information. Whenever you have more than one set of values, and a relationship exists between values in the separate sets, you may use a table to align the related values by placing them in the same row or column. For instance, tables are often used to display meeting agendas, with start times in one column, the names of the topics that will be covered in the next, and the names of the facilitators in the next, as in the following example:

Time	Topic	Facilitator
09:00 AM	Opening remarks	Scott Wiley
09:15 AM	Product demo	Sheila Prescott
10:00 AM	Discussion	Jerry Snyder
10:45 AM	Planning	Pamela Smart

FIGURE 3.8 This is an example of a table that does not contain quantitative data, in this case a meeting agenda.

Tables have been in use for almost two millennia, so they are readily understood by almost everyone who can read.

When to Use Tables

A handful of conditions should direct you to select a table, rather than a graph, as the appropriate means of display, but tables provide one primary benefit:

> Tables make it easy to look up individual values.

Tables excel as way to display simple relationships between quantitative values and the categorical items to which they're related so that these individual values can be easily located.

When deciding whether to use a table or graph to communicate your quantitative message, always ask yourself how the information will be used. Will you or others use it to look up one or more particular values? If so, it's a prime candidate for expression in a table. Or might the information be used to examine a set of quantitative values as a whole to discern patterns? If so, it's a prime candidate for expression in a graph, as we'll see soon.

Tables also make it easy to compare pairs of related values (e.g., sales in quarter 1 to sales in quarter 2). Here's a typical example:

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Monthly Average
2000	138.1	138.6	139.3	139.5	139.7	140.2	140.5	140.9	141.3	141.8	142.0	141.9	**140.3**
2001	142.6	143.1	143.6	144.0	144.2	144.4	144.4	144.8	145.1	145.7	145.8	145.8	**144.5**
2002	146.2	146.7	147.2	147.4	147.5	148.0	148.4	149.0	149.4	149.5	149.7	149.7	**148.2**
2003	150.3	150.9	151.4	151.9	152.2	152.5	152.5	152.9	153.2	153.7	153.6	153.5	**152.4**
2004	154.4	154.9	155.7	156.3	156.6	156.7	157.0	157.3	157.8	158.3	158.6	158.6	**156.9**
2005	159.1	159.6	160.0	160.2	160.1	160.3	160.5	160.8	161.2	161.6	161.5	161.3	**160.5**
2006	161.6	161.9	162.2	162.5	162.8	163.0	163.2	163.4	163.6	164.0	164.0	163.9	**163.0**
2007	164.3	164.5	165.0	166.2	166.2	166.2	166.7	167.1	167.9	168.2	168.3	168.3	**166.6**
2008	168.8	169.8	171.2	171.3	171.5	172.4	172.8	172.8	173.7	174.0	174.1	174.0	**172.2**
2009	175.1	175.8	176.2	176.9	177.7	178.0	177.5	177.5	178.3	177.7	177.4	176.7	**177.1**
2010	177.1	177.8	178.8	179.8	179.8	179.9	180.1	180.7	181.0	181.3	181.3	180.9	**179.9**

FIGURE 3.9 This is an example of a simple table that can be used to look up several years of monthly rates.

Tables work well for look-up and one-to-one comparisons, in part because their structure is so simple, and in part because the quantitative values are encoded as text, which we can understand directly, without translation. Graphs, by contrast, are visually encoded, which requires translation of the information into the numbers it represents.

The textual encoding of tables also offers a level of precision that cannot be provided by graphs. It is easy to express a number with as much specificity as you wish using text (e.g., 27.387483), but the visual encoding of individual numbers in graphs doesn't lend itself to such precision.

Another strength of tables is that they can include multiple sets of quantitative values that are expressed in different units of measure. For instance, if you need to provide sales information that includes the number of units sold, the dollar amount, and a comparison to a forecast expressed as a percentage, doing so in a single graph would be difficult, because a graph usually contains a single quantitative scale with a single unit of measure.

And a final strength of tables is their ability to combine summary and detail information in a single display. For example, a table might include the amount of revenue earned per month (detail), with a total for the year (summary).

To summarize, tables are used to display simple relationships between quantitative values and corresponding categorical items, which makes tables ideal for looking up and comparing individual values. The entire set of reasons to use a table consists of the following; if one or more of these are true, you should probably display the data in a table:

1. The display will be used to look up individual values.
2. It will be used to compare individual values but not entire series of values to one another.
3. Precise values are required.
4. The quantitative information to be communicated involves more than one unit of measure.
5. Both summary and detail values are included.

Graphs Defined

Graphs exhibit the following characteristics:

- Values are displayed within an area delineated by one or more axes.
- Values are encoded as visual objects positioned in relation to the axes.
- Axes provide scales (quantitative and categorical) that are used to label and assign values to the visual objects.

Axes delineate the space that is used to display data in a graph.

Essentially, a graph is a *visual display of quantitative information*. Whereas tables encode quantitative values as text, graphs encode quantitative values visually. Consider this simple example:

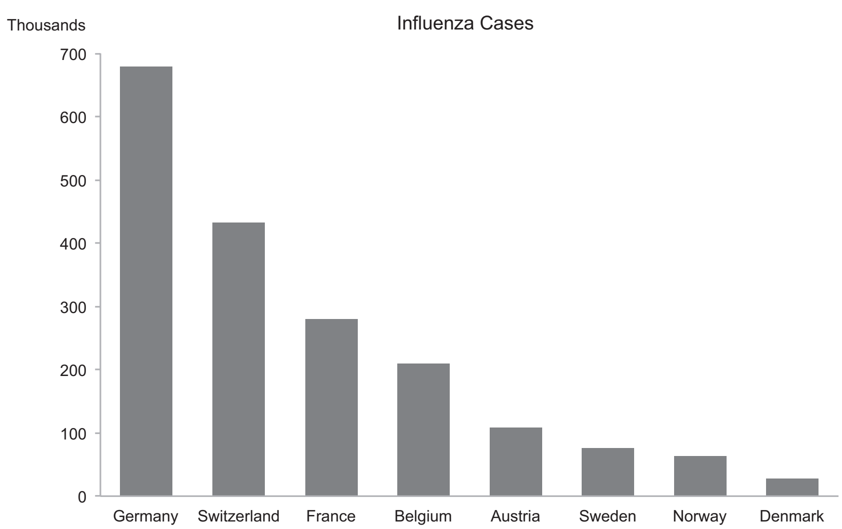

FIGURE 3.10 This is an example of a simple graph, which displays a count of influenza cases per country.

This graph has two axes: one that runs horizontally, called the *X axis*, and one that runs vertically, called the *Y axis*. In this graph, the categorical scale, which labels the countries, resides along the X axis, and the quantitative scale (i.e., counts of influenza cases) resides along the Y axis. The values themselves are encoded as rectangles, called *bars*. Bars are one of several visual objects that can be used to encode data in graphs. The number of influenza cases in each country is encoded as the height of its bar and the position of its top in relation to the scale on the Y axis. The horizontal position of each bar along the X axis is labeled to denote the specific categorical item to which the values are related (e.g., Belgium).

With a little practice, even someone who has never previously used a graph can learn to interpret the information contained in a simple one like this. Although the values of sales for each country cannot be interpreted to the exact number, this isn't the graph's purpose. Rather, the graph paints the picture that influenza has affected these countries to varying degrees, with the greatest effect in Germany and the least in Denmark, and significant differences in between. Information like this, which is intended to show patterns in data, is best presented in a graph rather than a table, as we'll see in the section *When to Use Graphs*. But first, a little history.

A Brief History of Graphs

Graphs of quantitative information have been in use for only a few hundred years, which is a relatively short time given the thousands of years that mathematics has existed. Despite how natural it may seem to see quantitative information displayed in graphs, the original notion that numbers could be displayed visually in relation to two perpendicular axes involved a leap of imagination. The launching pad for this leap had already been around for many centuries before quantitative graphs emerged. An earlier type of visual information display, also assisted by a scale of measurement along perpendicular axes, eventually suggested the possibility of graphs. Can you guess what it was? It is still in common use today for the purpose of navigation. It is a two-dimensional representation of the physical world that's used to measure distances between locations. I'm referring to a *map*. The earliest known map dates back about 4,300 years. It was drawn on a clay tablet and represented northern Mesopotamia. When a map depicts the entire world, or some large part of it, the standard set of grid lines that allow us to determine location and distance are longitudes and latitudes. These grid lines form a quantitative scale of sorts.

In the 17th century these grids were adapted for the representation of numbers alone. In his *La Géométrie* (1637), René Descartes introduced them as a means to encode numbers as coordinate positions in a two-dimensional (2-D) grid. This innovation provided the groundwork for an entire new field of mathematics that is based on graphs. To some extent even prior to Descartes, others had already experimented with graphical displays of numbers. Ron Rensink writes:

> *Although Descartes did contribute to the graphic display of quantitative data in the 17th century, other forms of quantitative graphs had already been used to represent things such as temperature and light intensity three centuries earlier. Indeed, as Manfredo Massironi discusses in his book (Massironi, M. (2002) The Psychology of Graphic Images: Seeing, Drawing, Communicating. Erlbaum, page 131), quantities such as displacement were graphed as a function of time as far back as the 11th century. But while these facts may be of interest in their own right, the more important point*

is that techniques in graphic representation have been developed over many centuries, and many of these techniques have been subsequently forgotten—perhaps fallen out of vogue, or never found wide use to begin with. But the reasons for their dismissal may not necessary apply in this day and age. Indeed, several techniques might lend themselves quite well to modern technology, and so might be worth resurrecting in one form or other. Books such as Massironi's are helpful in discovering such possibilities. [1]

Despite these earlier ventures, it wasn't until the late 18th century that the use of graphs to present numbers became popular. William Playfair, a Scottish social scientist, used his imagination and design acumen to invent many of the graphing techniques that we use today. He pioneered the use of graphs to reveal the shape of quantitative information, thus providing a way to communicate quantitative relationships that numbers expressed as text could never convey. He invented the bar graph, was perhaps the first person to use a line to show how values change through time, and on a day when he was probably under the weather, invented the pie chart as well, but we'll forgive his brief lapse of judgment.

1. Ron Rensink (2011) "Four Futures and a History," www.Interaction-Design.org.

Information about the early development of the graph, including its precursor, the map, may be found in Robert E. Horn (1998) *Visual Language*. MacroVU, Inc. Robert Horn provides an informative timeline, which cites the milestones in the historical development of visual information display.

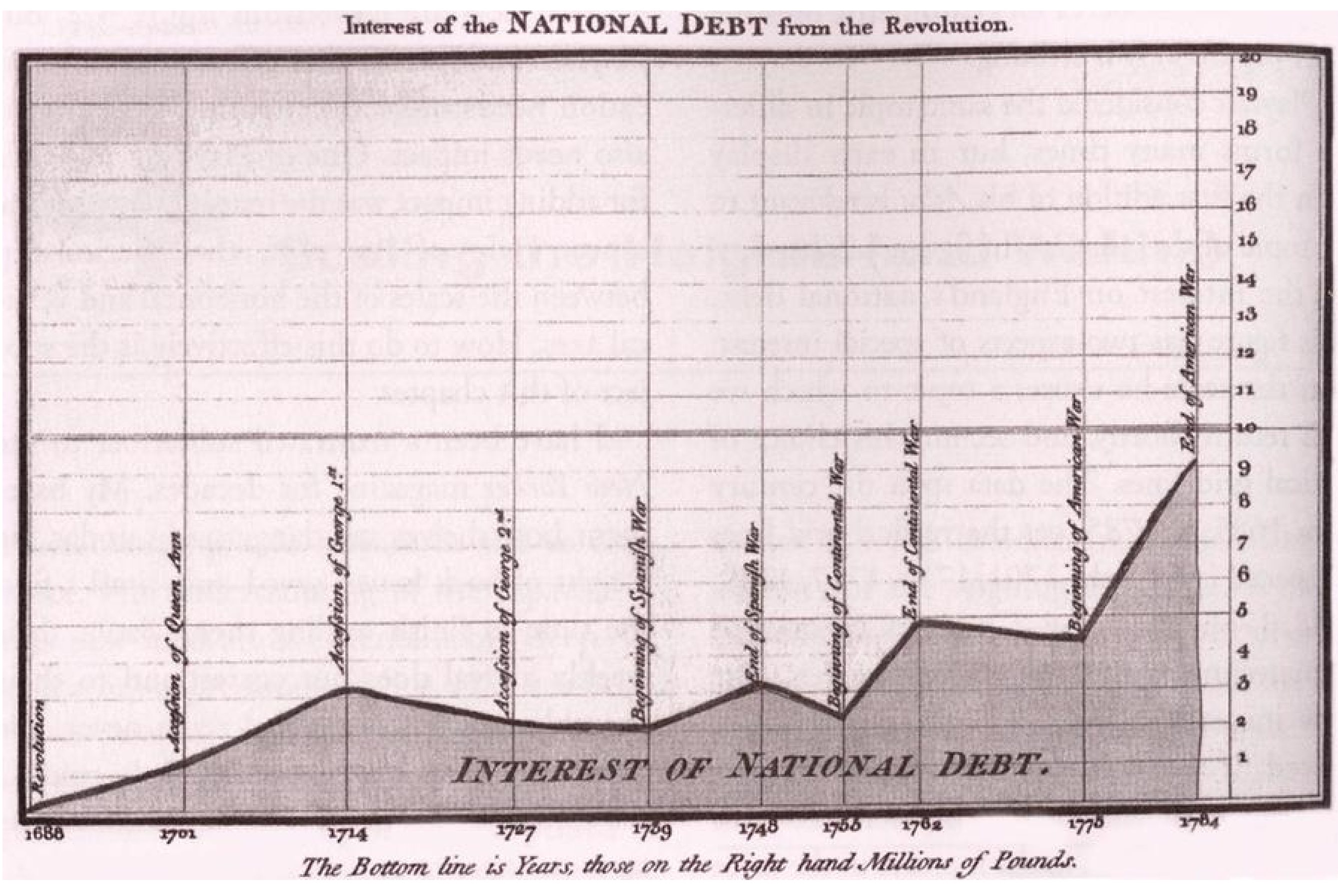

FIGURE 3.11 This graph was included in William Playfair's book *The Commercial and Political Atlas* in 1786 to make a case against England's policy of financing colonial wars through national debt.

The old saying, "A picture is worth a thousand words" applies quite literally to graphs. By presenting quantitative information in visual form, graphs efficiently communicate what might otherwise require a thousand or even a million words, and sometimes communicate what words could never convey.

From the time of Playfair until today, many innovators have added to the inventory of graph designs available for the representation, exploration, analysis, and communication of quantitative information. During the past 50 years, none has contributed more to the field as an advocate of excellence in graphic design than Edward Tufte, who in 1983 published his landmark treatise on the subject, *The Visual Display of Quantitative Information*. With the publication of additional books and articles since, Tufte continues to be respected as a major authority in this field.

The work of William S. Cleveland, especially his book *The Elements of Graphing Data*, is also an outstanding resource.[2] Cleveland's work is particularly useful to those with statistical training who are interested in sophisticated graphs, such as those used in scientific research.

2. William S. Cleveland (1994) *The Elements of Graphing Data.* Hobart Press.

When to Use Graphs

Graphs reveal more than a collection of individual values. Because of their visual nature, graphs present the overall shape of the data. Text, displayed in tables, cannot do this. The patterns revealed by graphs enable readers to detect many points of interest in a single collection of information. Take a look at the next example, and try to identify some of the features that graphs can reveal. Approach this by first determining the messages in the data, then by identifying the visual cues—the shapes—that reveal each of these messages.

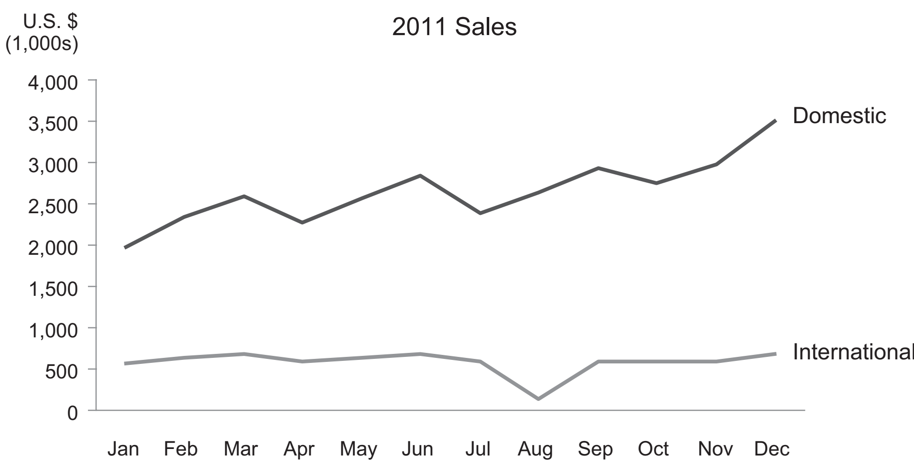

FIGURE 3.12 This fairly typical line graph is an example of the shapes and patterns in quantitative data that graphs make visible.

What insights are brought to your attention by the shape of this information? Take a moment to list them in the right margin.

• • • • • • •

Let's walk through a few of these revelations together, beginning with domestic sales. What does the shape of the black line tell you about domestic sales during the year 2011? One message is that, during the year as a whole, domestic sales increased; these sales ended higher than they started, with a gradual increase through most of the year. One name for this type of pattern is a *trend*, which displays the overall nature of change during a particular period of time.

Was this upward trend steady? No, it exhibits a pattern that salespeople sometimes call the *hockey stick*. Sales go down in the first month of each quarter and then gradually go up to a peak in the last month of each quarter. If you examine the shape of the line from the last month of one quarter, such as March, to the last month of the next quarter, such as June, you'll recognize that these segments each look a little like a hockey stick. This pattern in sales usually occurs when salespeople are given bonuses based on reaching or exceeding a quarterly quota.

Now, if you look at sales in the first quarter of the year versus sales in the second quarter, then the third, and finally the fourth, you find that the graph makes it easy to see how these different periods compare. When information is given shape in a graph, it becomes possible to compare entire sets of data. In this case the comparison is between different quarters.

Now look for a moment at international sales. Once again you are able to easily see the trend of sales throughout the year, which in this case is relatively flat. Compared with domestic sales, international sales seem to exhibit less fluctuation through time and relatively little difference between the beginning and end of the year. Although international sales appear to be less affected by quarters and seasons, one point along the line stands out as quite different from the rest: the month of August. Sales in August took an uncharacteristic dip compared to the rest of the year. As an analyst, this abnormal sales value would make you want to dig for the cause. If you did, perhaps you would discover that most of your international customers were vacationing in August and therefore weren't around to place orders. Whatever the cause, the current point of interest to us is that graphs make exceptions to general patterns stand out clearly from the rest. This wouldn't be nearly so obvious in a table of numbers.

Finally, if you widen your perspective to all the quantitative information, you find that the graph makes it easy to see the similarities and differences between the two sets of values (domestic and international sales), both overall and at particular points in the graph. We could go on, adding more to the list of characteristics that graphs reveal, but we've hit the high points and are now ready to distill what we've detected to its essence:

> Graphs are used to display relationships among and between sets of quantitative values by giving them shape.

The visual nature of graphs endows them with their unique power to reveal patterns of various types, including changes, differences, similarities, and exceptions. Graphs can communicate quantitative relationships that are much more complex than the simple associations between individual quantitative values that tables can express.

Graphs can display large data sets in a way that can be readily perceived and understood. You could gather data regarding the relationship between employee productivity and the use of two competing software packages, involving thousands of records across several years, and, with the help of a graph, you would be able to immediately see the nature of the relationship. If you have ever tried to use a huge table of data for analysis, you would quickly fall in love with graphs like scatter plots, which can make the relationships among thousands of individual data points instantly intelligible.

Patterns in data are difficult to discern even in a small table of numbers. Despite our agreement on most matters, Tufte and I disagree slightly about this. According to Tufte:

> *Tables usually outperform graphics in reporting on small sets of 20 numbers or less. The special power of graphics comes in the display of large data sets.* [3]

3. Edward Tufte (2001) *The Visual Display of Quantitative Information*, Second Edition. Graphics Press, page 56.

Although it is true that one's chance of seeing a meaningful pattern in small table of numbers is greater than in one that is large, those patterns are much easier to see in a graph. Take a look at the following table.

Job Satisfaction By Income, Education, and Age

	College Degrees		No College Degrees	
Income	Under 50	50 & over	Under 50	50 & over
Up to $50,000	643	793	590	724
Over $50,000	735	928	863	662

FIGURE 3.13 This table displays measures of employee job satisfaction, divided into categories.

Do any unusual patterns of job satisfaction pop out? Can you see that one group of employees exhibits a different pattern of satisfaction from the others? Now take a look at the following graph.

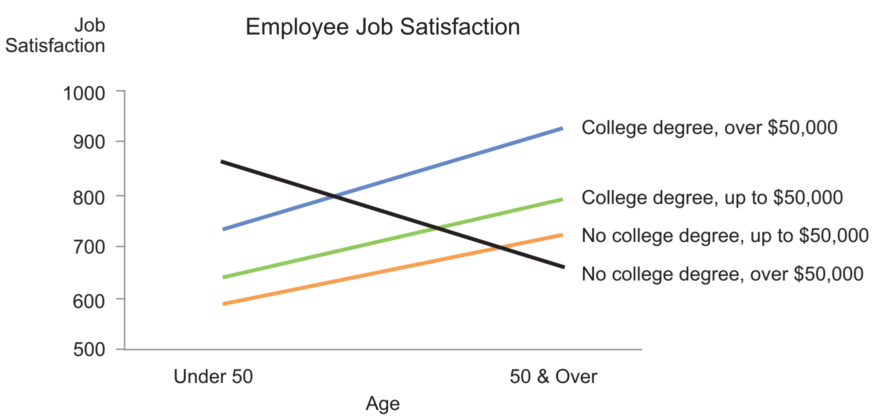

FIGURE 3.14 This graph causes the pattern represented by the black line to pop out.

The fact that employees with salaries over $50,000 but no college degrees experience a significant decrease in job satisfaction in their older years now jumps out. Even with only eight numbers, this graph did what the table couldn't do—it made this diverging pattern obvious. This results not only because the information is displayed graphically, but also because the graph was specifically designed to feature this particular pattern. Had I designed the graph in the following manner, that pattern would no longer be obvious.

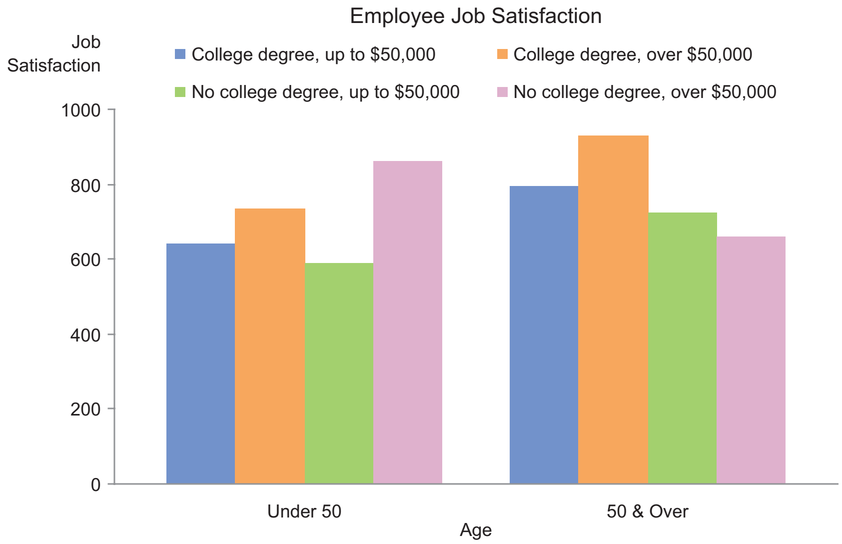

FIGURE 3.15 Unlike the graph in *Figure 3.14*, this graph does not make the pattern of decreasing job satisfaction among the one group of employees obvious.

This example is not based on real data, so if you make more than $50,000 a year but don't have a college degree and will soon turn 50, don't worry, you'll probably do just fine.

So displaying information in a graph rather than a table does not by itself make meaningful patterns visible. We must design the graph to feature evidence of the particular story that we're trying to tell. Many possible stories dwell in a set of data, even in a small one such as illustrated above. Later in this book, we'll learn to make the necessary graph design choices to tell particular stories.

Summary at a Glance

Use Tables When	Use Graphs When
• The display will be used to look up individual values. • It will be used to compare individual values. • Precise values are required. • The quantitative values include more than one unit of measure. • Both detail and summary values are included.	• The message is contained in the shape of the values (e.g., patterns, trends, and exceptions). • The display will be used to reveal relationships among whole sets of values.