Graphic Discovery

A Trout in the Milk and Other Visual Adventures

HOWARD WAINER

Princeton University Press Princeton and Oxford

Published by Princeton University Press,
41 William Street, Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
3 Market Place, Woodstock, Oxfordshire OX20 ISY

Second printing, and first paperback printing, 2008
Paperback ISBN: 978-0-691-13405-5

The Library of Congress has cataloged the cloth edition of this book as follows

Wainer, Howard.
Graphic Discovery : a trout in the milk and other visual
adventures / Howard Wainer.
p. cm.
Includes bibliographical references and indexes.
ISBN 0-691-10301-1 (cloth : acid-free paper)
1. Mathematical statistics—Graphic methods.
2. Playfair, William, 1759–1823.
3. Mathematical statistics—History—18th century. I. Title.
QA276.3.W29 2005
519.5—dc22 2003066083

British Library Cataloging-in-Publication Data is available

This book has been composed in
Adobe Garamond and Akzidenz Grotesk
Printed on acid-free paper. ∞
press.princeton.edu
Printed in the United States of America
10 9 8 7 6 5 4 3 2

Frontispiece: A portrait of Sir Francis Galton (1822–1911) by Susan Slyman showing him among some of his graphic inventions. In the upper left corner is the graph of his discovery of the anticyclonic movement of air around low-pressure zones that forms the basis of chapter 8.

Introduction

Let me begin with a few kind words about the bubonic plague. In 1538, Thomas Cromwell, the Earl of Essex (1485–1540),* issued an injunction (one of seventeen) in the name of Henry VIII that required the registration of all christenings and burials in every English parish. The London Company of Parish Clerks compiled weekly *Bills of Mortality* from such registers. This record of burials provided a way to monitor the incidence of plague within the city. Initially, these *Bills* were circulated only to government officials, principal among them the Lord Mayor and members of the King's Council.† They were first made available to the public in 1594, but were discontinued a year later with the abatement of the plague. However, in 1603, when the plague again struck London, their publication resumed on a regular basis.

The first serious analysis of the London *Bills* was done by John Graunt in 1662, but in 1710, Dr. John Arbuthnot, a physician to Queen Anne, published an article that used the christening data to support an argument (probably tongue-in-cheek) for the existence of God. These data also provide supporting evidence for the lack of existence of graphs at that time.

Figure 1 is a simple plot of the annual number of christenings in London from 1630 until 1710. The preparation of such a plot is straightforward, certainly requiring no more complex apparatus than was available to Dr. Arbuthnot in 1710. Moreover, as we will see in a moment, it is quite informative. Yet it is highly unlikely that Arbuthnot, or any of his contemporaries, ever made such a plot.

The overall pattern we see in figure 1 is a trend over these eighty years of an increasing number of christenings, almost doubling from 1630 to 1710. A number of fits and starts manifest themselves in substantial jiggles. Yet each jiggle, save one, can be explained. Some of these explanations are written on the plot. The big dip that began in 1642 can only partially be explained by the onset of the English Civil War. Surely the chaos common to civil war can explain the initial drop, but the war ended in 1649 with the beheading of Charles I at Whitehall, whereas the christenings did not return to their earlier levels until 1660.‡ Graunt offered a more complex explanation that involved the distinction between births and christenings, and the likelihood that Anglican priests would not enter children

* Thomas Cromwell was an English statesman who had a successful career as an administrator and advisor to the king. Among his other accomplishments, he arranged Henry VIII's divorce from Catherine of Aragon (and was largely responsible for the beheading of Sir Thomas More [1478–1535] in the process). Ironically, five years later, he was done in by Henry's aversion to Anne of Cleves, a consort of Cromwell's choosing, when he was consequently sent to the Tower of London and beheaded. Perhaps specifying the consequences of failure in this way would provide a workable pathway toward helping to improve the efficacy of computer dating services.

† This exposition is heavily indebted to the scholarly work of Sandy Zabell, to whose work the interested reader is referred for a much fuller description (Zabell 1976). It was Zabell who first uncovered Arbuthnot's clerical error.

‡ The year 1660 marked the end of the protectorate of Oliver Cromwell and the beginning of the restoration.

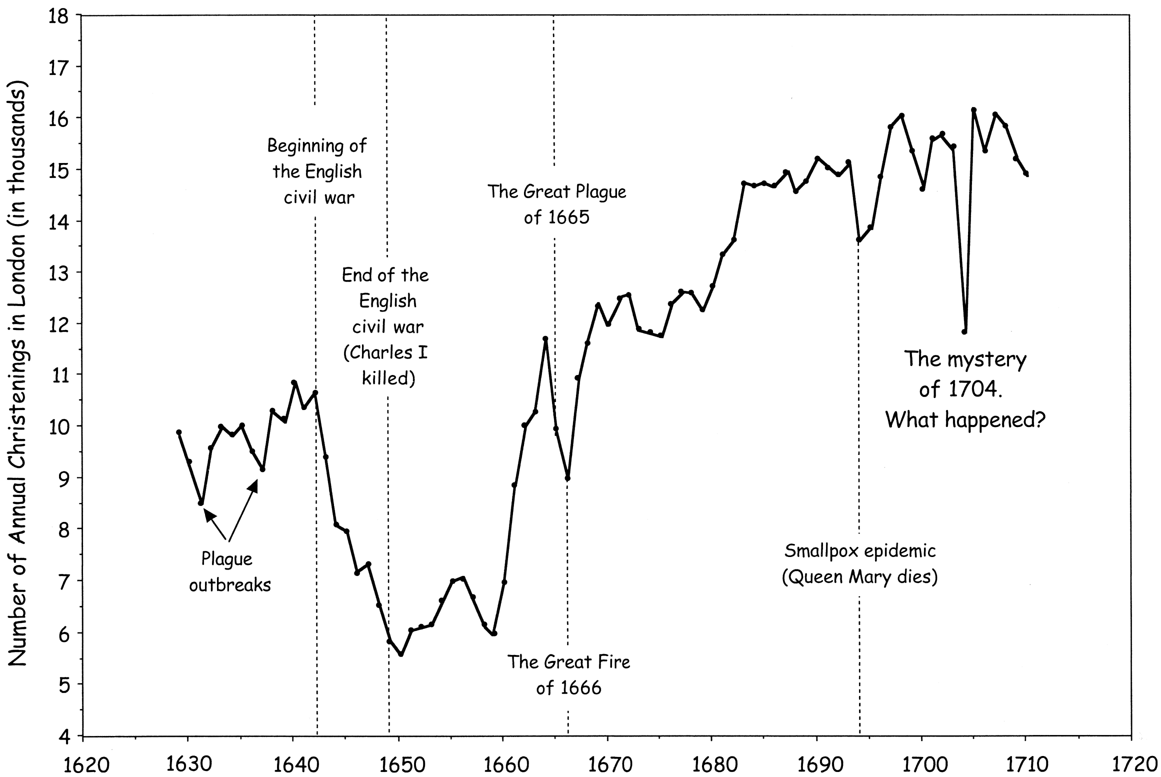

Figure 1. A plot of the annual christenings in London between 1630 and 1710 from the London *Bills of Mortality*. These data were taken from a table published by John Arbuthnot in 1710.

born to Catholics or Protestant dissenters into the register.

Many of the other irregularities observed are explained in figure 1, but what about the mysterious drop in 1704? That year has about four thousand fewer christenings than one might expect from observing the adjacent data points. What happened? There was no sudden outbreak of war or pestilence, no great civil uprising, nothing that could explain this enormous drop.

The plot not only reveals the anomaly, it also presents a credible explanation. In figure 2, I have duplicated the christening data and drawn a horizontal line across the plot through the 1704 data point. In doing so we immediately see that the line goes through exactly one other point—1674. If we went back to Arbuthnot's table we would see that in 1674 the number of christenings of boys and girls were 6,113 and 5,738, exactly the same number as he had for 1704. Thus the 1704 anomaly is likely to be a copying error! In fact, the correct figure for that year is 15,895 (8,153 boys and 7,742 girls), which lies comfortably between the christenings of 1703 and 1705 as expected.

It seems reasonable to assume that Arbuthnot, upon seeing such an unusual data point, would have investigated and, finding a clerical error, would have corrected it. Yet he did not. He did not, despite the fact that when graphed the error stood out, literally, like a sore thumb. Thus we must conclude that he never graphed his data. Why not?

The remarkable answer to this question occupies most of part I of this book. In brief,

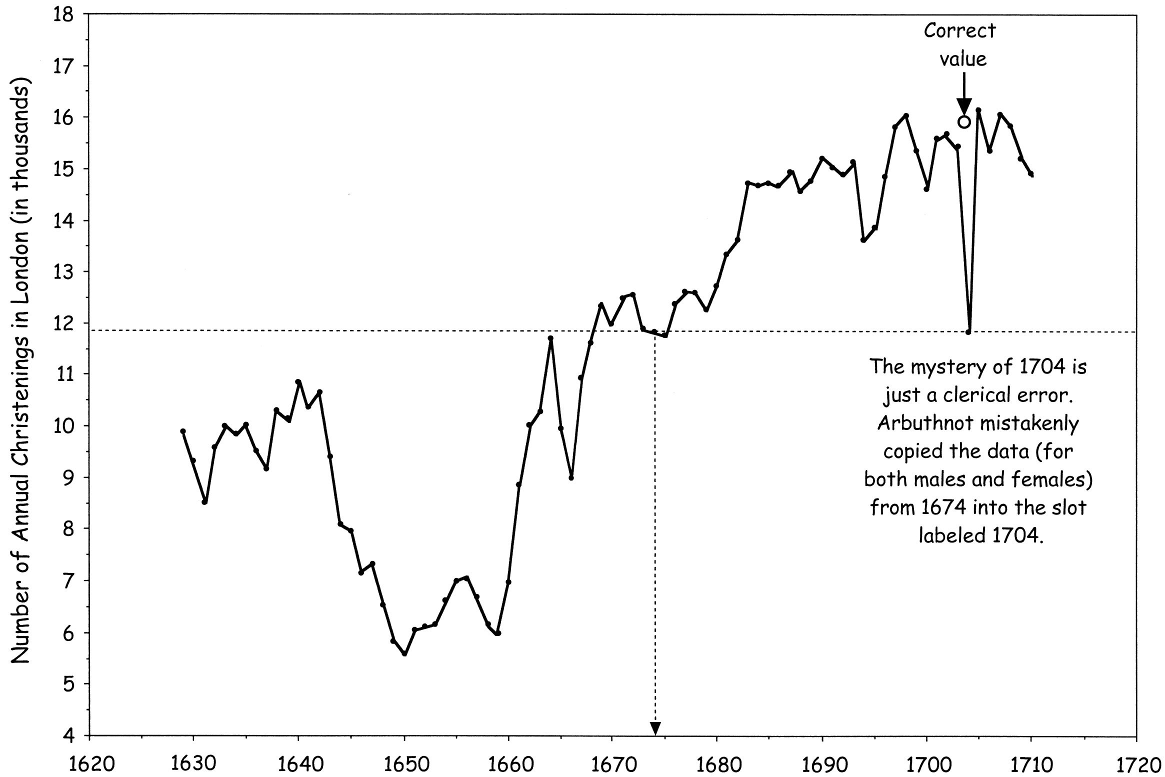

Figure 2. The solution to the mystery of 1704 is suggested by noting that only one other point (1674) had exactly the same values as the 1704 outlier. This coincidence provided the hint that allowed Zabell to track down Arbuthnot's clerical error.

it is that the very idea of graphing data was not yet invented. The story of its invention and inventor is a fascinating one that mixes science and politics, intrigue and scandal, revolution and shopping. The principal actors in this drama include some of the leading scientists of the day as well as a scoundrel of the first order. It also includes a French physician, a scientist of great renown, two signers of the American Declaration of Independence, and America's third president.

Part II illustrates the power of this marvelous invention to help us understand the modern world. It contains examples as disparate as a murder trial in Connecticut, the effect of the Vietnam war on college admissions, faxing policies in Canada, preposterous sports cars, the Boston Marathon, and the perambulations of the stock market over the past century. In each of these situations we see how a graphic depiction of data can help us see things that otherwise would have been as invisible to us as Arbuthnot's clerical error was to him.

Part III of this book looks to the future. I begin with a brief examination of the life and work of John Wilder Tukey, one of the twentieth century's great geniuses. I then describe some of the sorts of graphical methods, closely tied to modern computing, that are likely to become the bedrock of the way that we will be able to understand the torrent of data that has become the standard in our information-laden society.

Superficially, this is a book celebrating data display. But that phrase has two parts: *data* and *display*. Fine display formats are worthless without important data, and so in celebrating the marvelous invention of data-based graphics we must also save some kudos for those who undertook the often thankless task of gathering the data.* Where would the Johannes Keplers of history have been without the Tycho Brahes?

Graphics evolved because (i) there was a growing recognition that important questions could be answered with data,† (ii) data were being gathered to aid in the quest for such answers, and (iii) graphs were the best way to find both the structure and the surprises hidden in data. The London *Bills of Mortality* were one of the earliest organized efforts to gather extensive longitudinal data in the hope of aiding the government in its search for an effective policy for dealing with public catastrophes such as the bubonic plague. Thus, let me muster a weak thank-you to the plague for providing the immediate motivation for the sort of data gathering that has been the backbone of public health policy in the four hundred years since.

* It is critical, in our efforts to gain a glimpse of the cosmos, that we do not forget the importance of the demanding and sometimes boring intellectual regimen without which discovery is impossible. The famous conductor Pierre Montreux, rich in years and reputation, was once told by a young conductor in one of his master classes at Tanglewood that he, the student, was desperately seeking the "meaning" of the "ineffable essence of Mozart." Montreux congratulated him on his high aim, and then said it would do him no harm, meanwhile, to "learn how to keep the beat."

† In the mid-eighteenth century, Samuel Johnson pointed this out when he said, "To count is modern practice, the ancient method was to guess," but he was far from the first to recognize this, as Seneca too was aware of the difference: "Magnum esse solem philosophus probabit, quantus sit mathematicus." ("The philosopher says the sun is large, but the mathematician measures it." *Epistulae* 88.27)

I William Playfair and the Origins of Graphical Display

The graphic explosion of the nineteenth century that manifested itself in the publication of atlases in all aspects of the observational sciences had its origin in the intellectual turbulence of the eighteenth century. The hundred-year span between 1750 and 1850 saw a shift in the language of science from words to pictures.*

This shift began with the historical time charts of Jacques Barbeau-Dubourg (see chapter 7), which were followed closely by similar efforts from the Scottish philosopher Adam Ferguson and twelve years later from Joseph Priestley (chapter 5). Indeed, this idea proved so useful that Thomas Jefferson even used it to keep track of the price of vegetables in the Washington market (chapter 6). The scientific use of graphic displays had its origin with the Dutch polymath Christiaan Huygens (1629–1693), who developed the first survival chart in 1669, and accelerated when Martin Lister provided graphical summaries of weather data before the Oxford Philosophical Society on March 10, 1683 (chapter 1), which was quickly followed by many others who wanted to understand better the weather data that were now available with the invention of the barometer. All of the plots of these worthy scholars were precursors to the work of the Scot William Playfair (1759–1823), whose *Commercial and Political Atlas of England and Wales* contained no maps but instead beautifully polished versions of most of the common graphical forms in use today. This atlas, in which modern techniques of the graphical presentation of quantitative phenomena emerged fully developed, earned Playfair my nomination for the title of father of modern graphical display. (See chapters 1–4 for more about Playfair's contributions and his remarkable life.) The traditions of graphical display and weather were brought into a wonderful consilience by Francis Galton, who mapped weather reports of December 1861, drawn from all over western Europe, to show for the first time phenomena that had previously only been the hinted at (chapter 8). With Galton, the case for a transition to graphical representation of scientific phenomena was complete.

But others before Galton had already been converted. In 1878, the French physiologist Etienne Marey understood the value of graph-

* After the perfection of engraving in the fifteenth century, scientific documents often included lots of pictures, including abstract diagrams. What did not appear until after Playfair (with a few exceptions such as musical notation and those examples I give in chapter 1) was the diagramming of variates from empirical observations that went beyond the scale translations of space and time as movement in space (Biderman 1990).

ical representation. His graphic schedule of all the trains between Paris and Lyons (see figure I.1) provides a powerful illustration of the breadth of value of this approach. And, on the off chance that someone might have missed the point, he provided an explicit conclusion: "There is no doubt that graphical expression will soon replace all others whenever one has at hand a movement or change of state—in a word, any phenomenon. Born before science, language is often inappropriate to express exact measures or definite relations."[1] Marey was also giving voice to the movement away from the sorts of subjectivity that had characterized prior science in support of the more modern drive toward objectivity.* Although some cried out for the "insights of dialectic," "the power of arguments," "the insinuations of elegance," and the "flowers of language," their protestations were lost on Marey, who dreamed of a wordless science that spoke instead in high-speed photographs and mechanically generated curves, in images that were, as he put it, in the "language of the phenomena themselves."[2]

Historians have pointed out that "Let nature speak for itself" was the watchword of the new brand of scientific objectivity that emerged at the end of the nineteenth century. "At issue was not only accuracy but morality as well: the all-too-human scientists must, as a matter of duty, restrain themselves from imposing their hopes, expectations, generalizations, aesthetics, and even their ordinary language on the image of nature."[3] Mechanically produced graphic images would take over when human discipline failed. Marey and his contemporaries turned to mechanically produced images to eliminate human intervention between nature and representation. "They enlisted polygraphs, photographs, and a host of other devices in a near-fanatical effort to produce atlases—the bibles of the observational sciences"—documenting birds, fossils, human bodies, elementary particles, flowers, and economic and social trends that were certified free of human interference.[4]

The problem for nineteenth-century atlas makers was not a mismatch between world and mind, as it had been for seventeenth-century epistemologists, but rather a struggle with inward temptation. The moral remedies sought were those of self-restraint: images mechanically reproduced and published, warts and all; texts so laconic that they threatened to disappear entirely. Seventeenth-century epistemology aspired to the viewpoint of angels; nineteenth-century objectivity aspired to the self-discipline of saints. The precise observations and measurements of nineteenth-century science required taut concentration endlessly repeated. It was a vision of scientific work that glorified the plodding reliability of the bourgeois rather than the moody brilliance of the genius.†

The graphical representation of scientific phenomena served two purposes. Its primary function was standardizing phenomena in visual form, but it also served the cause of publicity for the scientific community. It preserved what was ephemeral and distributed it to all who would purchase the volume, not just the lucky few who were in the right place at the right time with the right

* Marey's view that graphical display could avoid the problems of language echoed the insight of his ill-fated countryman, Louis XVI, who expressed this view after receiving and reading a copy of Playfair's *Atlas* (see chapter 2).

† Although with such contributors as Condorcet (1743–1794), von Humboldt (1769–1859), and Florence Nightingale (1820–1910), there was certainly room for genius in the eighteenth and nineteenth centuries. Indeed Galton's weather maps, developed in the middle of the nineteenth, show how plodding reliability, when adjoined with moody brilliance, can yield especially fruitful results (chapter 8), yet no one would doubt that Robert Plot was a plodding plotter.

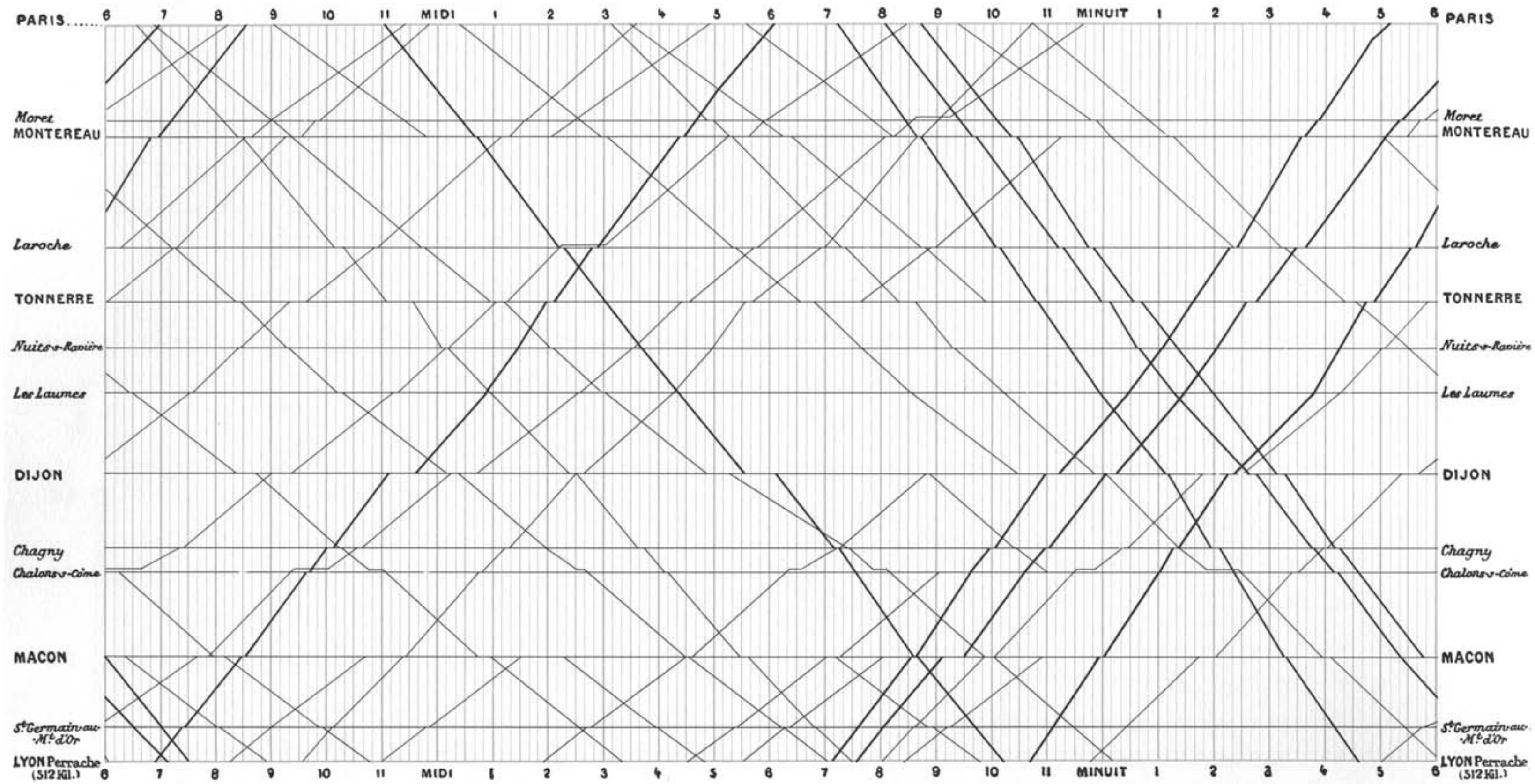

Figure I.1. Etienne Marey's (1878) "graphical train schedule" showing the daily passage of all of the trains between Paris and Lyons. This version, prepared by Edward Tufte (1983), uses a gray grid and is reproduced with his permission.

equipment. And it served the cause of memory, for images are more vivid and indelible than words.

But the graphical display of natural phenomena was viewed as yet more. Marey, in an accompanying note to his design of a portable polygraph, which automatically registered a variety of measures, suggested that through the use of graphics scientists could reform the very essence of scientific research and scientific evidence. "The graphic method translates all these changes in activity of forces into an arresting form that one could call the language of the phenomena themselves, as it is superior to all other modes of expression" (p. iv). Such a language was, for Marey, universal in two senses. Graphical representation could cut across the artificial boundaries of natural languages to reveal nature to all people, and it could cut across disciplinary boundaries to capture phenomena as diverse as the pulse of a heart and the downturn of an economy. Pictures became more than merely helpful tools: they were the words of nature herself.*

* This simple? Perhaps not. An alternative thesis to the one that characterizes science's task as capturing the glorious revelations by nature of her sublime design is one that sees man imposing the order of his senses and his arts upon the unheavenly disorder amidst which he finds himself.

1 Why Playfair?

"Getting information from a table is like extracting sunbeams from a cucumber" (Farquhar and Farquhar).* This evocative indictment of data tables by two nineteenth-century economists comes as no great insight to anyone who has ever tried to draw inferences from such a data display. For most purposes we almost always prefer a graphical representation. Indeed, graphs are ubiquitous now; hence it is hard to imagine a world before they existed. Yet data graphs are a human invention, indeed a relatively modern one. Data-based graphics began to make an appearance in the mid-seventeenth century but their full value and great popularity can be traced to a single event and a single person. The event was the publication, in 1786, of a small atlas describing the imports and exports of England and Wales with their various trading partners. The atlas contained forty-four graphs and no maps. Its author was a twenty-seven-year-old Scot named William Playfair, and his *Commercial and Political Atlas* forever changed the way that we look at data.

William Playfair (1759–1823) worked as a draftsman for James Watt and was the ne'er-do-well younger brother of the well-known scientist John Playfair (1748–1814).† William Playfair is often credited with being the progenitor of modern statistical graphics. Most histories of statistical graphics give him huge credit while acknowledging important graphical work that preceded him.[1] A balanced summary is that he invented many of the currently popular graphical forms,‡ improved the few that already existed, and broadly popularized the idea of graphic depiction of quantitative information. Before Playfair, statistical graphics were narrowly employed and even more narrowly circulated. After him, graphs popped up everywhere, being used to convey information in the social, physical, and natural sciences.

The title comes from Albert Biderman's private characterization of the question that has intrigued him for more than a decade. The intellectual contents of this chapter come principally from two sources: Biderman's 1978 talk at the Leesburg, Virginia, conference on Social Graphics (and its published elaboration, Biderman 1990) and Patricia Costigan-Eaves and Michael Macdonald-Ross's extensive, but as yet unpublished, history of early graphic developments. Some of their material is in Costigan-Eaves and Macdonald-Ross 1990.

* This well-known quotation, though pithy, is somewhat inaccurate. What the brothers Farquhar actually wrote (1891, p. 55) was, "The graphical method has considerable superiority for the exposition of statistical facts over the tabular. A heavy bank of figures is grievously wearisome to the eye, and the popular mind is as incapable of drawing any useful lessons from it as of extracting sunbeams from cucumbers."

† John Playfair's activities were remarkably varied: minister, geologist, mathematician, and professor of natural philosophy at Edinburgh University. In fact, in 1805 William thanked his brother for the idea of using "lines applied to matters of finance" that William used in his 1786 book. We can only speculate why it took him nineteen years to give his brother some credit.

‡ *Invented* is perhaps too strong a term. It may be more accurate to refer to him as an important deployer of graphical forms. He did not invent so much as he permuted and manip-

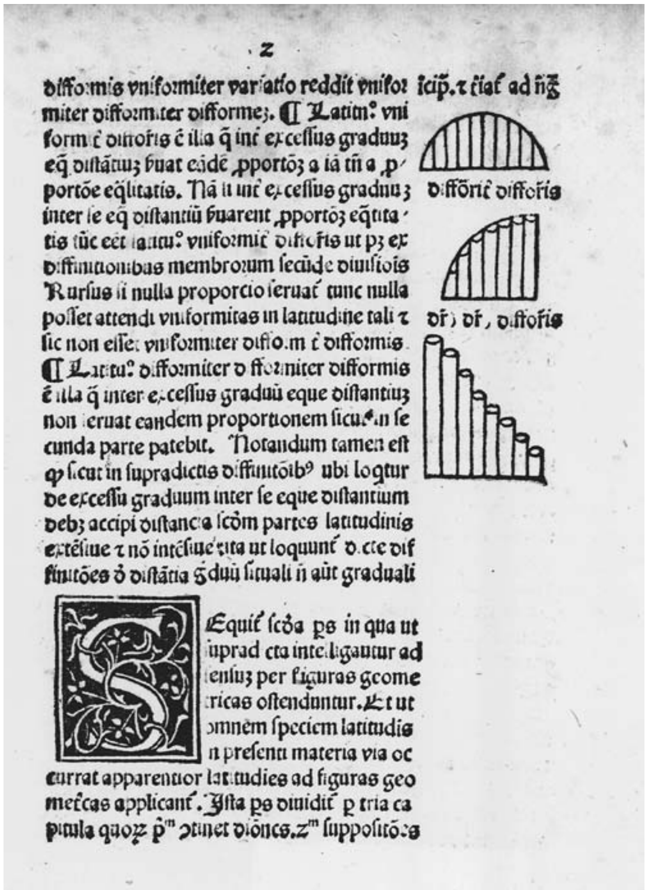

Figure 1.1. The first page of Oresme's *Tractatus de latitudunibus formarum*, the Padua edition of 1486. This item is reproduced by permission of the Huntington Library, San Marino, California.

Before we meet Playfair, however, it is worthwhile to step back a century or so and examine the attitudes that pervaded scientific investigations. Because natural science originated within natural philosophy, it favored a rational rather than an empirical approach to scientific inquiry. Such an outlook was antithetical to the more empirical modern approach to science, which does not disdain the atheoretical plotting of data points with the goal of investigating suggestive patterns. Graphs that were in existence before Playfair (with some notable exceptions that I will discuss shortly) grew out of the same rationalist tradition that yielded Descartes's coordinate geometry—that is, the plotting of curves on the basis of an a priori mathematical expression. For example, Oresme's "pipes" on the first page of the Padua edition of his 1486 *Tractatus de latitudunibus formarum* (figure 1.1) is often cited as an early example.*

This notion is supported by statements such as that of Luke Howard, a prolific grapher of data in the late eighteenth and early nineteenth centuries who, as late as 1844, apologized for his methodology and referred to it as an "autograph of the curve . . . confessedly adapted rather to the use of the *dilettanti* in natural philosophy than that of regular students."

All the mechanical pieces necessary for data-based graphics were in place long before Playfair. For example, a primitive coordinate system of intersecting horizontal and vertical lines that enable a precise placement of data points was used by surveyors of the Nile flood basin as early as 1400 B.C. A more refined coordinate system was used by Hipparchus (ca. 140 B.C.), whose terms for the coordinate axes translate into Latin as *longitudo* and

---

ulated graphic elements; he varied graphic codes and formats as experiments to try to improve the visual portrayal of quantitative relationships. But there is no doubt that Playfair's use of graphics was more influential than his predecessors'. Part of this must have been because Playfair's graphs were so beautifully produced (compare his line charts with those of Huygens), but more important is the undeniable fact that Playfair published statistical graphics for all to see. Moreover, he did this repeatedly, and with a coherent theme, thus powerfully making the point that the graphical depiction of information can communicate quantitative information in an accurate and relatively painless way.

* Clagett (1968) argued convincingly that this work was not written by Oresme, but probably by Jacobus de Sancto Martino, one of his followers, in about 1390—yet another instance of how surprisingly often eponymous referencing is an indication only of who did not do it (Stigler 1980).

latitudo, to locate points in the heavens. Somewhat later, Roman surveyors used a coordinate grid to lay out their towns on a plane that was defined by two axes. The *decimani* were lines running from east to west, and the *cardi* ran north to south.[2] Many other special-purpose coordinate systems were in wide use before Playfair: for example, musical notation placed on horizontal running lines was in use as early as the ninth century,[3] and the chessboard was invented in seventh-century India.

One of the earliest examples of printed graph paper dates from about 1680.[4] Large sheets of paper engraved with a grid were apparently printed to aid in designing and communicating the shapes of the hulls of ships.* Many historians describe Descartes's 1637 development of a coordinate system as an important intellectual milestone in the path toward statistical graphics.[5] More recent work interprets this in exactly the opposite way—as an intellectual impediment that took a century and a half and Playfair's eclectic mind to overcome.[6]

Although the use of coordinate grids is very old, graphic encoding of information is older still. Paleolithic cave art provides an early and very striking example of graphic display. Some Ice Age bone carvings of animals are intermixed with patterns of dots and strokes that some archeologists have interpreted as a lunar notation system related to the animals' seasonal appearance. These are almost identical in structure, as well as degree of detail, to the engraving on the hull of the Pioneer 10 spacecraft that shows a drawing of a man and a woman along with a simple plotting of the Earth's location by dotted pulsar beams.

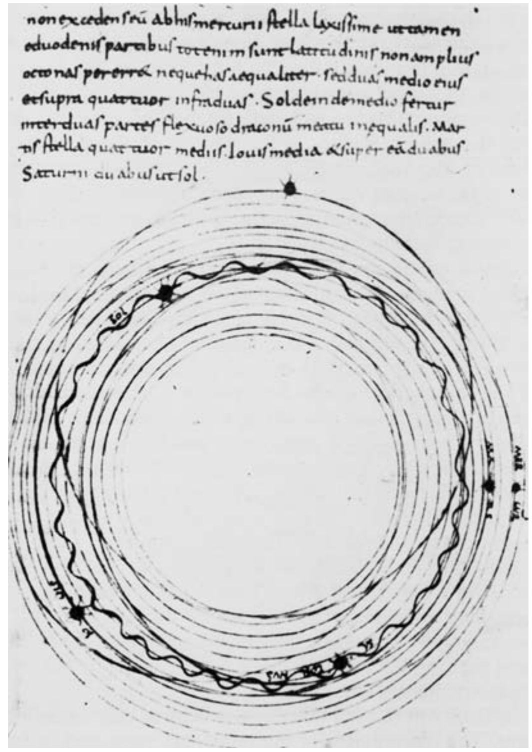

Figure 1.2. Plinian circular diagram of planetary latitudes, early ninth century. Taken from Eastwood (1987), p. 158, figure 6.

So we have the ideas of graphic encoding of information and a coordinate grid system. Why not the plotting of data? Well, some data were plotted. Let us consider three examples.

Example 1. Pliny's Ninth-Century Astronomical Charts

Pliny's (ca. 810) astronomical data were plotted in a roughly circular form (see figure 1.2) corresponding to the varying locations of the bodies in the heavens. But these graphs (astronomical maps) were not as useful as they might have been, because locating a body in the circular path required visually tracking the

* This material is classed in the "collection" category of the British Library with the entry, "A collection of engraved sheets of squared paper, whereon are traced in pencil or ink the curves or sweeps of the hulls of sundry men-of-war."

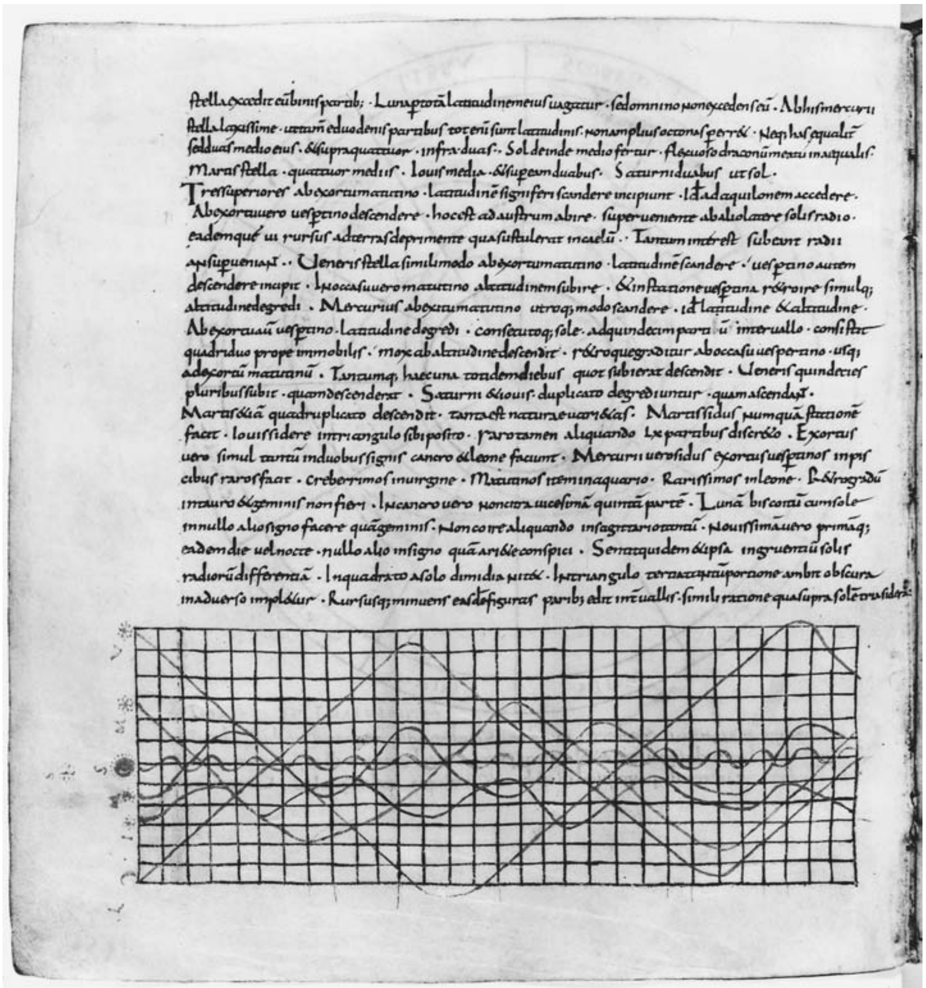

Figure 1.3. Plinian planetary latitudes, portrayed on a rectangular grid. Photo courtesy Burgerbibliothek Bern, ms. 367 f. 24v.

complete cycle. A manuscript originating in Auxerre toward the end of the ninth century contained a scheme that remedied this by transforming the circular grid into a rectangular one (see figure 1.3). The cyclic nature of the orbit is less apparent, but by making explicit the time (horizontal axis) and height above the horizon (vertical axis) it made locating and identifying a heavenly body somewhat easier.*

Example 2. Christiaan Huygens's Seventeenth-Century Survival Charts

On October 30, 1669, the Dutch polymath Christiaan Huygens (1629–1693) received a letter from his brother Lodewijk containing some interpolations of life expectancy data taken from John Graunt's 1662 book *Natural and Political Observations on the Bills of Mortality*. Christiaan responded in letters

* There is actually less here than meets the eye. The horizontal axis is not particularly well defined. It is really an unfolding of the circular version with the horizontal spacing being only roughly related to time

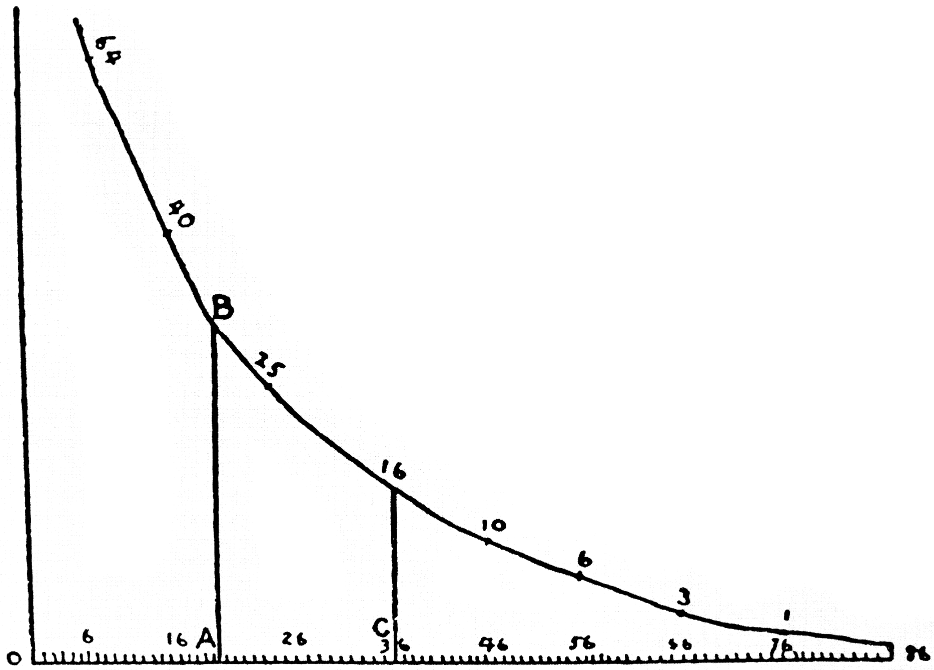

Figure 1.4. Christiaan Huygens's 1669 curve showing how many people out of a hundred survive between the ages of infancy and eighty-six. The data are taken from John Graunt, *Natural and Political Observations on the Bills of Mortality* (1662).

dated November 21 and 28, 1669, with graphs of those interpolations. Figure 1.4 contains one of those graphs showing age on the horizontal axis and number of survivors of the original birth cohort on the vertical axis. The curve drawn was fitted to his brother's interpolations.* The letters on the chart are related to an associated discussion on how to construct a life expectancy chart from this one—that is, analyzing a set of data to gain deeper insights into the subject. Christiaan constructed such a chart and indicated that it was more interesting from a scientific point of view; figure 1.4, he felt, was more helpful in wagering.†

Example 3. Robert Plot's Seventeenth-Century Plots of Barometric Pressure

Good graphs can make difficult problems trivial. We have all become used to weather forecasts that are very accurate and detailed for a

* Huygens's twenty-two-volume *Oeuvres complètes* (1888–1950) contains many other graphical devices to be explored by anyone with fluency in ancient Dutch, Latin, and French. Incidentally, Huygens's graphical work on the pendulum proved to him that a pendulum's oscillations would be isochronic regardless of its amplitude. This discovery led him to build the first clock based on this principle.

† This scooped a 1976 paper by the Chicago statistician Sandy Zabell, whose graphical analysis of the *Bills of Mortality* found inconsistencies, clerical errors, and a remarkable amount of other information, "much of it unappreciated at the time of their publication" (Zabell 1976, p. 27). Zabell's point, though implicit in his paper, is important in this discussion. As we illustrated in one situation in the introduction to this book, this was strong evidence that graphic display was not widely available. For had they been seen, these errors, which could not be missed with any sort of competent display, would have been discovered and eliminated.

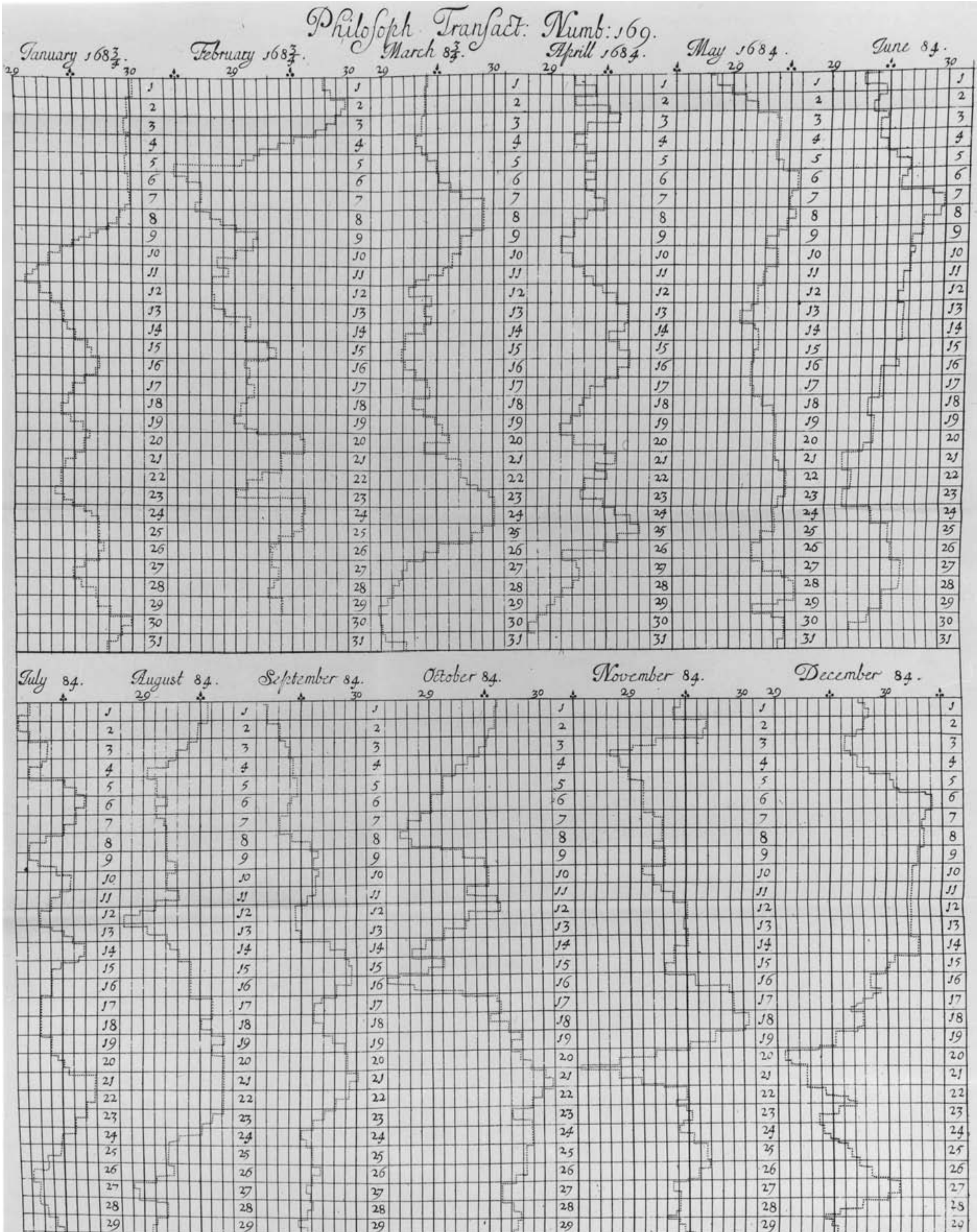

Figure 1.5. Robert Plot's (1685) "History of the Weather" recording of the daily barometric pressure in Oxford for the year 1684. Appears in *Philosophical Transactions* and is based on the original work of Martin Lister. Photo © The Royal Society.

day or two and pretty good for as far in advance as a week. I used to think that this was due to the increasing sophistication of complex prediction models.* But then I noticed the weather maps shown on every news broadcast. Using a model of no greater sophistication than that employed by Benjamin Franklin (weather generally moves from west to east), I was able to predict that the area of precipitation currently over Ohio would be hitting New Jersey by tomorrow and would stay over us until the weekend. Any fool could see it. The improvement in forecasting has not been entirely due to improvements in the mathematical models of the weather. The enormous wealth of radar and satellite data summarized into a multicolored and dynamic graphic can turn anyone into an expert.

The path to modern weather graphs is more than three hundred years long. The barometer was developed in 1665. Robert Plot recorded the barometric pressure in Oxford every day in 1684 and summarized his findings in a remarkably contemporary graph (figure 1.5) that he called a "History of the Weather."

He sent a copy of this graph with a letter to Martin Lister† in 1685 with a prophetic insight:

> For when once we have procured fit persons enough to make the same Observations in many foreign and remote parts, how the winds stood in each, at the same time, we shall then be enabled with some grounds to examine, not only the coastings, breadth, and bounds of the winds themselves, but of the weather they bring with them; and probably in time thereby learn, to be forewarned certainly, of divers emergencies (such as heats, colds, dearths, plague, and other epidemical distempers) which are not unaccountable to us; and by their causes be instructed for prevention, or remedies: thence too in time we may hope to be informed how far the positions of the planets in relation to one another, and to the fixed stars, are concerned in the alterations of the weather, and in bringing and preventing diseases and other calamities . . . we shall certainly obtain more real and useful knowledge in matters in a few years, than we have yet arrived to, in many centuries.

With so many predecessors, why have I chosen Playfair as my candidate for the father of modern graphical display? Although the arguments will build in subsequent chapters, a visual comparison of a sampling of Playfair's plots with any of those that came before him makes clear the qualitative jump that Playfair's work represents. Ian Spence points out that Playfair was the first to use hachure, color, and area to represent quantities in a systematic way. Moreover, Playfair published these forms in widely circulated volumes. Most of all, however, Playfair's graphs provided proof that the presentation of evidence could be beautiful.

As but one such example, compare Playfair's harmonious depiction of more than a century of data on England's national debt

* It is true that models are more sophisticated than they were in the past. I was enormously impressed when some surprising turns in a hurricane's path were predicted well in advance, but such models seem to be needed no more often than seldom.

† The origin of the graphical depiction of weather data, sadly, for the obvious eponymous glory, rests not with Plot but rather with Lister, who presented various versions of graphical summaries of weather data before the Oxford Philosophical Society on March 10, 1683, and later in the same year presented a modified version to the Royal Society. Plot was not the only one enthusiastic about Lister's graphical methods. William Molyneux was so taken that he had an engraving made of the grid, and he faithfully sent a "Weather Diary" monthly to William Musgrave. One of Molyneux's charts was reproduced in Gunther (1968).

(figure 1.6) with any of the graphs produced previously by others. The viewer's eye is drawn from the soaring debt to the vertical lines that communicate the events that presaged a change in the debt. The viewer's mind cannot avoid making the causal inference suggested. Nothing that had been produced before was even close. Even today, after more than two centuries of graphical experience, Playfair's graphs remain exemplary standards for clear communication of quantitative phenomena.

It now seems wise to recapitulate the argument: Graphical forms were available before Playfair, but they were rarely used to plot empirical information. I argue that this was because there was an antipathy toward the empirical approach. This suggestion is supported by statements such as that made by Luke Howard. But at least sometimes when data were available (for example, Pliny's astronomical data, Graunt's survival data, Plot's weather data, and several other admirable uses), they were plotted. Could it be that the exponential increase in the use of graphics after the publication of Playfair's *Atlas* was merely concomitant to the exponential growth in the availability of data? Or did the availability of graphic devices for analyzing data encourage data gathering? And why Playfair? Was he merely at the cusp of an explosion in data gathering and so his graphic efforts appear causal? Or did he play an important role in that explosion?

The consensus of scholars is that until Playfair "many of the graphic devices used were the result of a formal and highly deductive science. . . . This world view was more comfortable with an arm-chair, rationalistic approach to problem-solving which usually culminated in elegant mathematical principles" often associated with elegant geometrical diagrams.[7] The empirical approach to problem solving, a critical driving force for data collection, was slow to emerge. But the empirical approach began to demonstrate remarkable success in solving problems, and with improved communications,* the news of these successes and hence the popularity of the associated graphic tools began to spread quickly.

So the picture is almost complete. The fundamental tools for the graphical display of data were available; there was an increase in the acceptance of an empirical approach to science as an important part of the scientific process; data were being gathered in greater and greater quantities; and the success of empirical procedures in solving important practical problems was being more widely communicated. This explains the growth of the graphical method, but still leaves the initial question, "Why Playfair?"

We are accustomed to intellectual diffusion taking place from the natural and physical sciences into the social sciences; certainly that is the direction taken for both calculus and the scientific method. But statistical graphics in particular, and statistics in general, went the reverse route. Although, as we have seen, there were applications of data-based graphics in the natural sciences, only after Playfair applied them in the social sciences did their popularity begin to accelerate. Playfair should be credited with producing the first chartbook of social statistics; indeed, publishing an *Atlas* that contained not a single map is one indication of his belief in the methodology (to say nothing of his chutzpah). Playfair's work was immediately admired, but emulation, at least

* The first encyclopedia in English appeared in 1704. The number of scientific periodicals began a rapid expansion at the end of the eighteenth century; between 1780 and 1789 twenty new journals appeared, and between 1790 and 1800 twenty-five more (McKie 1972)

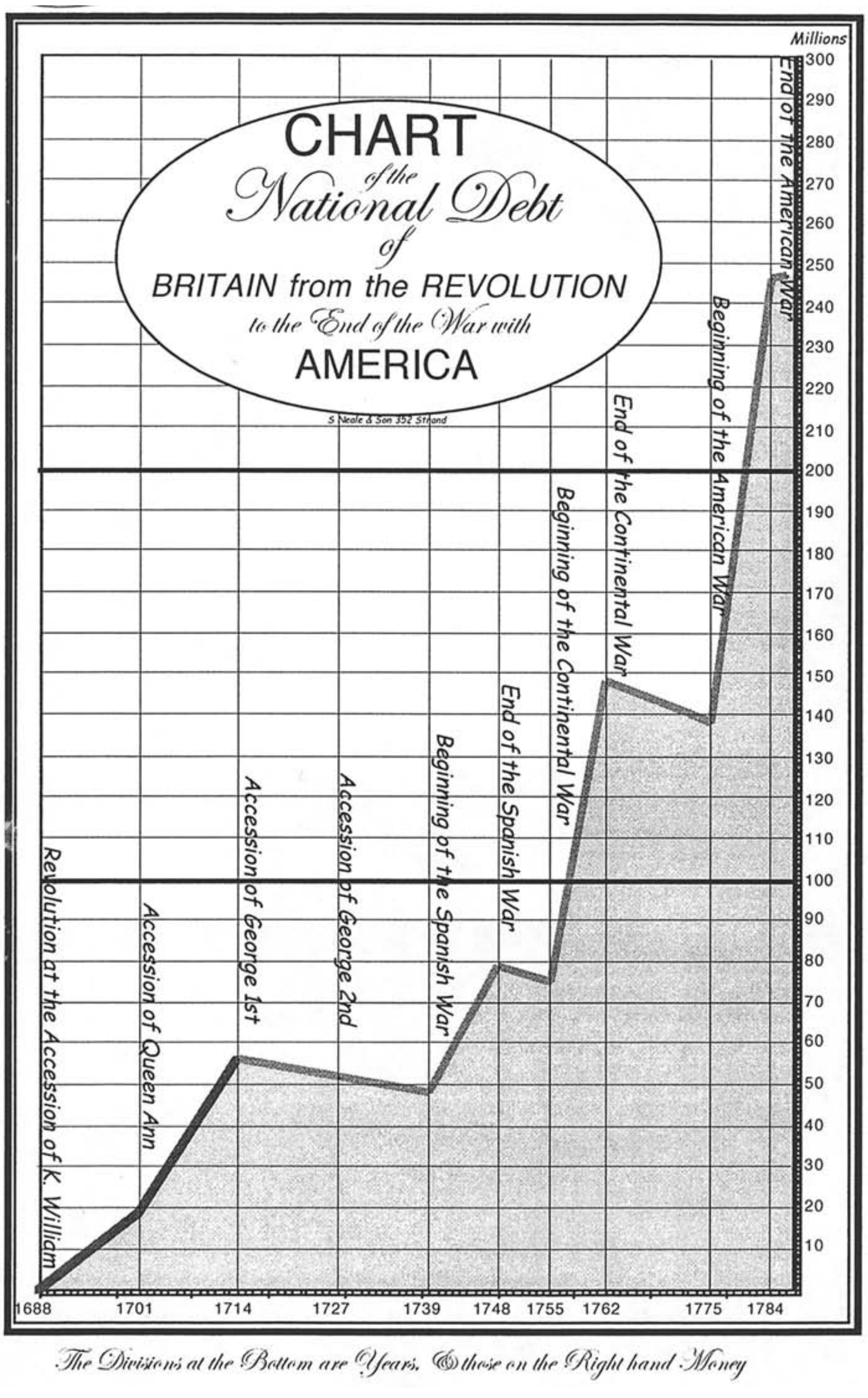

Figure 1.6. A close facsimile of William Playfair's plot of England's national debt from 1688 until 1786. It appeared in his *Commercial and Political Atlas* and accompanied his discussion arguing against the British government's policy of financing its colonial wars through debt.

in Britain, took a little longer (graphics use started on the continent a bit sooner). Interestingly, one of Playfair's earliest emulators was the banker S. Tertius Galton (the father of Francis Galton, and hence the biological grandfather of modern statistics), who in 1813 published a multiline time-series chart of the money in circulation, rates of foreign exchange, and prices of bullion and of wheat.* The relatively slower diffusion of the graphical method back into the natural sciences provides additional support for the hypothesized bias against empiricism there. The newer social sciences, having no such tradition and faced with both problems to solve and relevant data, were quicker to see the potential of Playfair's methods.

Playfair's graphical inventions and adaptations look contemporary. He invented the statistical bar chart out of desperation, because he lacked the time-series data required to draw a line showing the trade with Scotland and so used bars to symbolize the cross-sectional character of the data he did have. Playfair acknowledged Priestley's priority in this form, although Priestley used bars to symbolize the life spans of historical figures in a time line.[8] (See chapters 5, 6, and 7 for more on the fascinating history of time lines and graphical display of historical data.)†

Playfair's role was crucial for several reasons, but it was not for his development of the graphical recording of data; others preceded him in that. Indeed, in 1805 he pointed out that as a child his brother John had him keep a graphical record of temperature readings. But Playfair was in a remarkable position. Because of his close relationship with his brother and his connections with James Watt, he was on the periphery of science. He was close enough to know of the value of the graphical method, but sufficiently detached in his own interests to apply them in a very different arena—that of economics and finance.

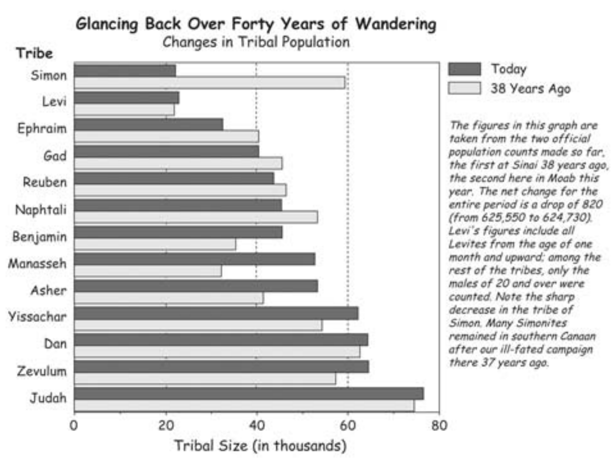

Figure 1.7. A translated and computer-enhanced reproduction of perhaps the earliest statistical graphic yet uncovered. It was apparently constructed about 1400 B.C. and was preserved in a sealed ceramic container in the Qumran caves. It was purchased by the author from Moishe the mapman at his Dead Sea antiquities stall in 1991.

* Biderman (1978) pointed out that, ironically, Galton's chart predicted the financial crisis of 1831 that created a ruinous run on his own bank.

† Priestley's use of the bar as a metaphor is somewhat different than Playfair's in that the data were not really statistical. A much earlier precedent has been recovered from its resting place in the Qumran caves abutting the Dead Sea. The graphic dates from approximately 1400 B.C. and was prepared as a summary of population changes in the twelve tribes of Israel as they emerged from their almost four decades of wandering in the Sinai after their exodus from Egypt, which began in April 1446 B.C. A faithful copy of this bar chart, with the captions and legends translated from Aramaic, is reproduced here as figure 1.7. Some aspects of this historic figure have been computer-enhanced for better reproduction. Note that it presages Huygens in subject matter and Playfair in form.

These areas, then as now, tend to attract a larger audience than matters of science, and Playfair was adept at self-promotion.

In a review of his 1786 *Atlas* that appeared in the *Political Herald*, Dr. Gilbert Stuart wrote,

> The new method in which accounts are stated in this work, has attracted very general notice. The propriety and expediency of all men, who have any interest in the nation, being acquainted with the general outlines, and the great facts relating to our commerce are unquestionable; and this is the most commodious, as well as accurate mode of effecting this object, that has hitherto been thought of. . . . To each of his charts the author has added observations [that] . . . in general are just and shrewd; and sometimes profound. . . . Very considerable applause is certainly due to this invention; as a new, distinct, and easy mode of conveying information to statesmen and merchants.

Such wholehearted approval rarely greets any scientific development. Playfair's adaptation of graphic methods to matters of general interest provided an enormous boost to the popularity of statistical graphics. His energy and artistic sense showed themselves in the forty color charts in his initial *Atlas*. The size of the undertaking required to produce such a book indicates Playfair's deep understanding of the power of the graphical method. His energy and skill as a draftsman, coupled with that understanding, led him to communicate his enthusiasm both widely and effectively. However, to be able to focus on graphics when the prevailing view of science looked upon such an approach as generally illegitimate requires a willingness to go against the tide—indeed, perhaps even taking joy in being an iconoclast. The events described in the next two chapters illuminate this aspect of Playfair's personality.

In *Kagemusha*, a film by the great Japanese director Akira Kurosawa, a legendary warlord is mortally wounded. The warlord's staff finds a petty thief, who bears a remarkable physical resemblance to the fallen leader, to substitute for him. With the substitute in place, the political strategy evolved by the dead warlord succeeds in his absence. In this examination of the question "the man or the moment?" Kurosawa clearly favors the latter. The Playfair enigma represents another instance of this great theme, although unlike Kurosawa's fictional situation, the more limited information available to us does not allow unambiguous conclusions.