Milestones in the History of Data Visualization: A Case Study in Statistical Historiography

Michael Friendly
Psychology Department and Statistical Consulting Service
York University
4700 Keele Street, Toronto, ON, Canada M3J 1P3

in: Classification: The Ubiquitous Challenge. See also BibTeX entry below.

BibTeX:

```
@incollection{Friendly:05:gfkl,
  author =       {M. Friendly},
  title =        {Milestones in the History of Data Visualization:
                  A Case Study in Statistical Historiography},
  year =         {2005},
  publisher =    {Springer},
  address =      {New York},
  booktitle =    {Classification:  The Ubiquitous Challenge},
  editor =       {C. Weihs and W. Gaul},
  pages =        {34--52},
  url =          {http://www.math.yorku.ca/SCS/Papers/gfkl.pdf},
}
```


YORK
UNIVERSITÉ
UNIVERSITY

Milestones in the History of Data Visualization: A Case Study in Statistical Historiography

Michael Friendly

Psychology Department
York University, Toronto, Canada
friendly@yorku.ca

Abstract. The Milestones Project is a comprehensive attempt to collect, document, illustrate, and interpret the historical developments leading to modern data visualization and visual thinking. This paper provides an overview and brief tour of the milestones content, with a few illustrations of significant contributions to the history of data visualization. This forms one basis for exploring interesting questions and problems in the use of statistical and graphical methods to explore this history, a topic that can be called "statistical historiography."

1 Introduction

The only new thing in the world is the history you don't know.—Harry S Truman

The graphic portrayal of quantitative information has deep roots. These roots reach into the histories of the earliest map-making and visual depiction, and later into thematic cartography, statistics and statistical graphics, medicine, and other fields, which are intertwined with each other. They also connect with the rise of statistical thinking and widespread data collection for planning and commerce up through the 19th century. Along the way, a variety of advancements contributed to the widespread use of data visualization today. These include technologies for drawing and reproducing images, advances in mathematics and statistics, and new developments in data collection, empirical observation and recording.

From above ground, we can see the current fruit; we must look below to understand their germination. Yet the great variety of roots and nutrients across these domains, that gave rise to the many branches we see today, are often not well known, and have never been assembled in a single garden, to be studied or admired.

The Milestones Project is designed to provide a broadly comprehensive and representative catalog of important developments in *all* fields related to the history of data visualization. Toward this end, a large collection of images, bibliographical references, cross-references and web links to commentaries on these innovations has been assembled.

This is a useful contribution in its own right, but is a step towards larger goals as well. First, we see this not as a static collection, but rather a dynamic database that will grow over time as additional sources and historical contributions are uncovered or suggested to us. Second, we envisage this project as providing a tool to enable researchers to work with or study this history, finding themes, antecedents, influences, patterns, trends, and so forth. Finally, as implied by our title, work on this project has suggested several interesting questions subsumed under the self-referential term "statistical historiography."

1.1 The *Milestones Project*

The past only exists insofar as it is present in the records of today. And what those records are is determined by what questions we ask.—Wheeler (1982, p. 24)

There are many historical accounts of developments within the fields of probability (Hald, 1990), statistics (Pearson, 1978, Porter, 1986, Stigler, 1986), astronomy (Riddell, 1980), cartography (Wallis and Robinson, 1987), which relate to, *inter alia*, some of the important developments contributing to modern data visualization. There are other, more specialized accounts, which focus on the early history of graphic recording (Hoff and Geddes, 1959, 1962), statistical graphs (Funkhouser, 1936, 1937, Royston, 1970, Tilling, 1975), fitting equations to empirical data (Farebrother, 1999), cartography (Friis, 1974, Kruskal, 1977) and thematic mapping (Palsky, 1996, Robinson, 1982), and so forth; Robinson (1982, Ch. 2) presents an excellent overview of some of the important scientific, intellectual, and technical developments of the 15^{th}–18^{th} centuries leading to thematic cartography and statistical thinking.

But there are no accounts that span the entire development of visual thinking and the visual representation of data, and which collate the contributions of disparate disciplines. In as much as their histories are intertwined, so too should be any telling of the development of data visualization. Another reason for interweaving these accounts is that practitioners in these fields today tend to be highly specialized, often unaware of related developments in areas outside their domain, much less their history. Extending Wheeler (1982), the records of history also exist insofar as they are collected, illustrated, and made coherent.

The initial step in portraying the history of data visualization was a simple chronological listing of milestone items with capsule descriptions, bibliographic references, markers for date, person, place, and links to portraits, images, related sources or more detailed commentaries. Its current public and visible form is that of hyper-linked, interactive documents available on the web and in PDF form (`http://www.math.yorku.ca/SCS/Gallery/milestone/`). We started with the developments listed by Beniger and Robyn (1978) and incorporated additional listings from Hankins (1999), Tufte (1983, 1990, 1997), Heiser (2000), and others. With assistance from *Les Chevaliers*, many other contributions, original sources, and images have been added. As explained below, our current goal is to turn this into a true multi-media database, which can be searched in flexible ways and can be treated as data for analysis.

2 Milestones Tour

In organizing this material, it proved useful to divide history into epochs, each of which turned out to be describable by coherent themes and labels. In the larger picture— recounting the history of data visualization— each milestone item has a story to be told: What motivated this development? What was the communication goal? How does it relate to other developments? What were the pre-cursors? What makes it a milestone? To illustrate, we present just a few exemplars from a few of these periods. For brevity, we exclude the earliest period (pre-17^{th} century) and the most recent period (1975–present) in this description.

2.1 1600-1699: Measurement and theory

Among the most important problems of the 17th century were those concerned with physical measurement— of time, distance, and space— for astronomy, surveying, map making, navigation and territorial expansion. This century also saw great new growth in theory and the dawn of practice— the rise of analytic geometry, theories of errors of measurement and estimation, the birth of probability theory, and the beginnings of demographic statistics and "political arithmetic."

As an example, Figure 1 shows a 1644 graphic by Michael Florent van Langren, a Flemish astronomer to the court of Spain, believed to be the first visual representation of statistical data (Tufte, 1997, p. 15). At that time, lack of a reliable means to determine longitude at sea hindered navigation and exploration.[1] This 1D line graph shows all 12 known estimates of the difference in longitude between Toledo and Rome, and the name of the astronomer (Mercator, Tycho Brahe, Ptolemy, etc.) who provided each observation.

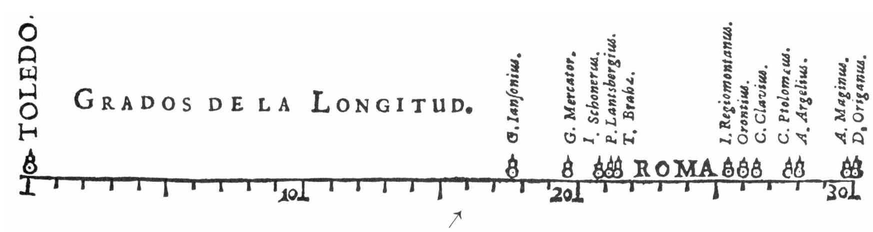

Fig. 1. Langren's 1644 graph of determinations of the distance, in longitude, from Toledo to Rome. The correct distance is $16°30'$. Source: Tufte (1997, p. 15)

What is notable is that van Langren could have presented this information in various tables— ordered by author to show provenance, by date to show priority, or by distance. However, only a graph shows the wide variation in the estimates; note that the range of values covers nearly half the length of the scale. Van Langren took as his overall summary the center of the range, where there happened to be a large enough gap for him to inscribe "ROMA." Unfortunately, all of the estimates were biased upwards; the true distance ($16°30'$) is shown by the arrow. Van Langren's

[1] For navigation, latitude could be fixed from star inclinations, but longitude required accurate measurement of time at sea, an unsolved problem until 1765.

graph is also a milestone as the earliest-known exemplar of the principle of "effect ordering for data display" (Friendly and Kwan, 2003).

2.2 1700-1799: New graphic forms

The 18th century witnessed, and participated in, the initial germination of the seeds of visualization that had been planted earlier. Map-makers began to try to show more than just geographical position on a map. As a result, new graphic forms (isolines and contours) were invented, and thematic mapping of physical quantities took root. Towards the end of this century, we see the first attempts at the thematic mapping of geologic, economic, and medical data.

Abstract graphs, and graphs of functions were introduced, along with the early beginnings of statistical theory (measurement error) and systematic collection of empirical data. As other (economic and political) data began to be collected, some novel visual forms were invented to portray them, so the data could "speak to the eyes."

As well, several technological innovations provided necessary nutrients. These facilitated the reproduction of data images (color printing, lithography), while other developments eased the task of creating them. Yet, most of these new graphic forms appeared in publications with limited circulation, unlikely to attract wide attention.

William Playfair (1759–1823) is widely considered the inventor of most of the graphical forms widely used today— first the line graph and bar chart (Playfair, 1786), later the pie chart and circle graph (Playfair, 1801). A somewhat later graph (Playfair, 1821), shown in Figure 2, exemplifies the best that Playfair had to offer with these graphic forms. Playfair used three parallel time series to show the price of wheat, weekly wages, and reigning monarch over a $\sim$250 year span from 1565 to 1820, and used this graph to argue that workers had become better off in the most recent years.

2.3 1800-1850: Beginnings of modern graphics

With the fertilization provided by the previous innovations of design and technique, the first half of the 19th century witnessed explosive growth in statistical graphics and thematic mapping, at a rate which would not be equalled until modern times.

In statistical graphics, all of the modern forms of data display were invented: bar and pie charts, histograms, line graphs and time-series plots, contour plots, scatterplots, and so forth. In thematic cartography, mapping progressed from single maps to comprehensive atlases, depicting data on a wide variety of topics (economic, social, moral, medical, physical, etc.), and introduced a wide range of novel forms of symbolism.

To illustrate this period, we choose an 1844 "tableau-graphique" (Figure 3) by Charles Joseph Minard, an early progenitor of the modern mosaic plot (Friendly, 1994). On the surface, mosaic plots descend from bar charts, but Minard introduced two simultaneous innovations: the use of divided and proportional-width bars so that area had a concrete visual interpretation. The graph shows the transportation

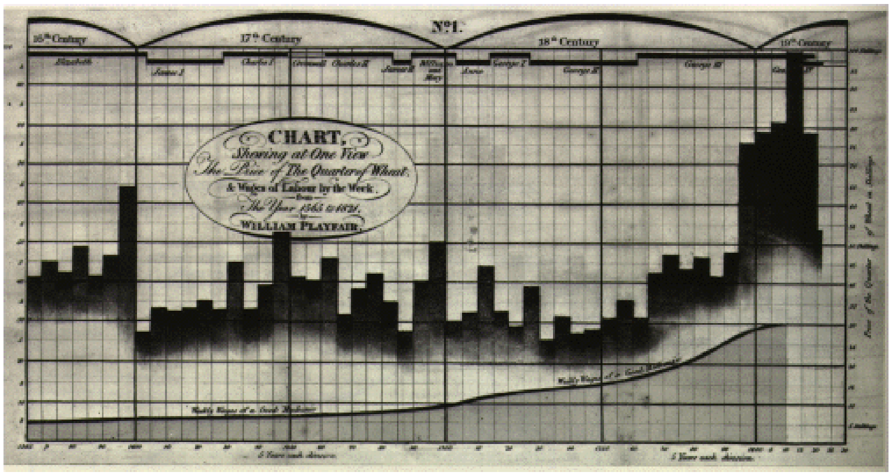

Fig. 2. William Playfair's 1821 time series graph of prices, wages, and ruling monarch over a 250 year period. *Source*: Playfair (1821), image from Tufte (1983, p. 34)

of commercial goods along one canal route in France by variable-width, divided bars (Minard, 1844). In this display the width of each vertical bar shows distance along this route; the divided bar segments have height $\sim$ amount of goods of various types (shown by shading), so the area of each rectangular segment is proportional to cost of transport. Minard, a true visual engineer (Friendly, 2000), developed such diagrams to argue visually for setting differential price rates for partial vs. complete runs. Playfair had tried to make data "speak to the eyes," but Minard wished to make them "calculer par l'oeil" as well.

2.4 1850-1900: The Golden Age of statistical graphics

By the mid-1800s, all the conditions for the rapid growth of visualization had been established. Official state statistical offices were established throughout Europe, in recognition of the growing importance of numerical information for social planning, industrialization, commerce, and transportation. Statistical theory, initiated by Gauss and Laplace, and extended to the social realm by Quetelet, provided the means to make sense of large bodies of data.

What started as the *Age of Enthusiasm* (Palsky, 1996) for graphics may also be called the *Golden Age*, with unparalleled beauty and many innovations in graphics and thematic cartography.

2.5 1900-1950: The modern dark ages

If the late 1800s were the "golden age" of statistical graphics and thematic cartography, the early 1900s could be called the "modern dark ages" of visualization (Friendly and Denis, 2000).

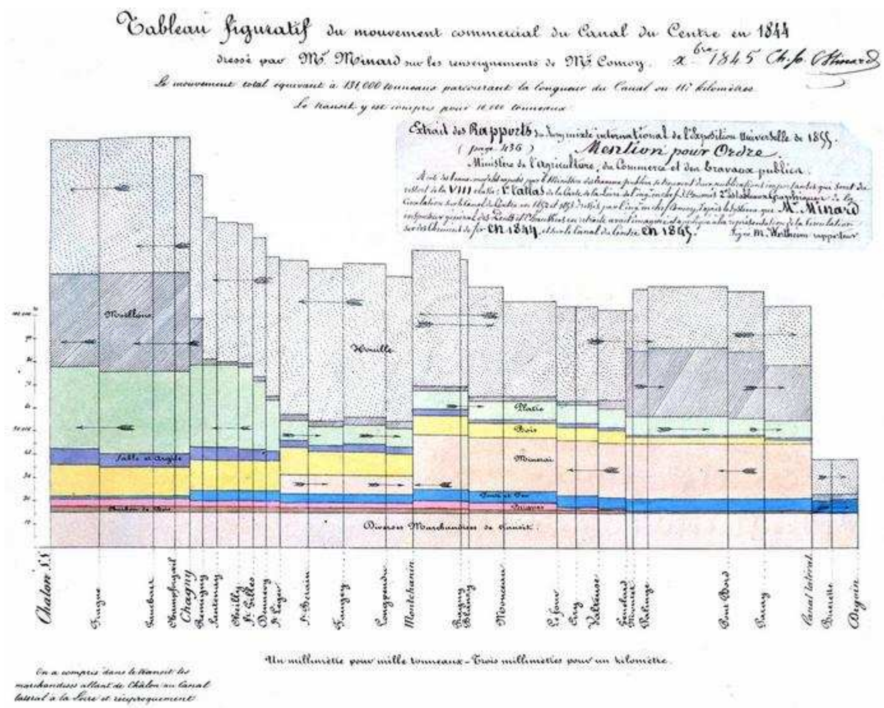

Fig. 3. Minard's Tableau Graphique, showing the transportation of commercial goods along the Canal du Centre (Chalon–Dijon). Intermediate stops are spaced by distance, and each bar is divided by type of goods, so the area of each tile represents the cost of transport. Arrows show the direction of transport. *Source:* ENPC:5860/C351 (Col. et cliché ENPC; used by permission)

There were few graphical innovations, and, by the mid-1930s, the enthusiasm for visualization which characterized the late 1800s had been supplanted by the rise of quantification and formal, often statistical, models in the social sciences. Numbers, parameter estimates, and, especially, standard errors were precise. Pictures were—well, just pictures: pretty or evocative, perhaps, but incapable of stating a "fact" to three or more decimals. Or so it seemed to statisticians.

But it is equally fair to view this as a time of necessary dormancy, application, and popularization, rather than one of innovation. In this period statistical graphics became main stream. It entered textbooks, the curriculum, and standard use in government, commerce and science. In particular, perhaps for the first time, graphical methods proved crucial in a number of scientific discoveries (e.g. the discovery of atomic number by Henry Mosely, lawful clusterings of stars based on brightness and color in the Hertzsprung-Russell diagrams; see Friendly and Denis (2004) for details.)

2.6 1950-1975: Re-birth of data visualization

Still under the influence of the formal and numerical zeitgeist from the mid-1930s on, data visualization began to rise from dormancy in the mid 1960s, spurred largely by three significant developments:

(a) In the USA, John W. Tukey began the invention of a wide variety of new, simple, and effective graphic displays, under the rubric of "Exploratory Data Analysis." (b) In France, Jacques Bertin published the monumental *Sémiologie Graphique* (Bertin, 1967, 1983). To some, this appeared to do for graphics what Mendeleev had done for the organization of the chemical elements, that is, to organize the visual and perceptual elements of graphics according to the features and relations in data. (c) Finally, computer processing of data had begun, and offered the possibility to construct old and new graphic forms by computer programs. True high-resolution graphics were developed, but would take a while to enter common use.

By the end of this period significant intersections and collaborations would begin: (a) computer science research (software tools, C language, UNIX, etc.) at Bell Laboratories (Becker, 1994) and elsewhere would combine forces with (b) developments in data analysis (EDA, psychometrics, etc.) and (c) display and input technology (pen plotters, graphic terminals, digitizer tablets, the mouse, etc.). These developments would provide new paradigms, languages and software packages for expressing statistical ideas and implementing data graphics. In turn, they would lead to an explosive growth in new visualization methods and techniques.

Other themes began to emerge, mostly as initial suggestions: (a) various visual representations of multivariate data (Andrews' plots, Chernoff faces, clustering and tree representations); (b) animations of a statistical process; and (c) perceptually-based theory (or just informed ideas) related to how graphic attributes and relations might be rendered to better convey the data visually.

3 Problems and Methods in Statistical Historiography

As we worked on assembling the Milestones collection, it became clear that there were several interesting questions and problems related to conducting historical research along these lines.

3.1 What counts as a Milestone?

In order to catalog the contributions to be considered as "milestones" in the history of data visualization, it is necessary to have some criteria for inclusion: for form, content, and substantive domain, as well as for "what counts" as a milestone in this context. We deal only with the last aspect here.

We have adopted the following scheme. First, we decided to consider several types of contributions as candidates: true innovations, important pre-cursors and developments or extensions. Second, we have classified these contributions according to several themes, categories and rubrics for inclusion. Attributions without reference here are listed in the Milestones Project web documents.

- ***Contributions to the development and use of graphic forms***. In statistical graphics, inventions of the bar chart, pie chart, line plot (all attributed to Playfair), the scatterplot (attributed to J.F.W. Herschel; see Friendly and Denis (2004)), 3D plots (Luigi Perozzo), boxplot (J. W. Tukey), and mosaic plot (Hartigan & Kleiner) provided new ways of representing statistical data. In thematic cartography, isolines (Edmund Halley), choropleths (Charles Dupin) and flow maps (Henry Harness; C. J. Minard) considerably extended the use of a map-based display to show more than simple geographical positions and features.
- ***Graphic content: data collection and recording***. Visual displays of information cannot be done without empirical data, so we must also include contributions to measurement (geodesy), recording devices, collection and dissemination of statistical data (e.g., vital statistics, census, social, economic data).
- ***Technology and enablement***. It is evident that many developments had technological prerequisites, and conversely that new technology allowed new advances that could not have been achieved before. These include advances in (a) reproduction of printed materials (printing press, lithography), (b) imaging (photography, motion pictures), and (c) rendering (computing, video display).
- ***Theory and practice***. Under this heading we include theoretical advances in the treatment and analysis such as (a) probability theory and notions of errors of measurement, (b) data summarization (estimation and modelling), (c) data exposure (e.g., EDA), as well as (d) awareness and use of these ideas and methods.
- ***Theory and data on perception of visual displays***. Graphic displays are designed to convey information to the human viewer, but how people use and understand this form of communication was not systematically studied until recent times. As well, proposals for graphical standards, and theoretical accounts of graphic elements and graphic forms provided a basis for thinking of and designing visual displays.
- ***Implementation and dissemination***. New techniques become *available* when they are introduced, but additional steps are needed to make them widely *accessible* and useable. We are thinking here mainly of implementations of graphical methods in software, but other contributions fall under this heading as well.

3.2 Who gets credit?

All of the Milestones items are attributed to specific individuals where we have reason to believe that names can be reasonably attached. Yet, *Stigler's Law of Eponomy* (Stigler, 1980) reminds us that standard attributions are often not those of priority. The Law in fact makes a stronger claim: “No scientific discovery is named after its original discoverer.” As *prima facie* evidence, Stigler attributes the origin of this law to Merton (1973).

As illustrations, Stigler (1980) states that Laplace first discovered the Fourier transform, Poisson first discovered the Cauchy distribution, and both de Moivre and Laplace have prior claims to the Gaussian distribution. He concludes that eponoyms are conveyed by the community of scholars, not by historians.

Thus, although all of the events listed are correctly attributed to their developers, it cannot be claimed with certainty that we are always identifying the first instance, nor that we give credit to all who have, perhaps independently, given rise to a new idea or method. Similarly, in recent times there may be some difficulty distinguishing credit among developers of (a) an underlying method or initial demonstration, (b) a corresponding algorithm, or (c) an available software implementation.

3.3 Dating milestones

In a similar way, there is some unavoidable uncertainty in the dates attached to milestone items, in a degree which generally increases as we go back in time. For example, in the 18^{th} and 19^{th} centuries, many papers were first read at scientific meetings, but recorded in print some years later; William Smith's geological map of England was first drawn in 1801, but only finished and published in 1815; some pre-1600 dates are only known approximately.

In textual accounts of history this does not present any problem— one can simply describe the circumstances and range of events, dated specifically or approximately, contributing to some development.

It does matter, however, if we wish to treat item dates as data, either for retrieval or analysis/display. For retrieval, we clearly want any date within a specified range to match; for analysis or display, the end points will sometimes be important, but sometimes it will suffice to use a middle value.

3.4 What is milestones "data"

The *Milestones Project* represents ongoing work. We continually update the web and pdf versions as we add items and images, many of which have been contributed by *Les Chevaliers*. To make this work, we rely on software tools to generate different versions from a single set of document sources, so that all versions can be updated automatically. For this, we chose to use LaTeX and BibTeX.

More recently, we have developed tools to translate this material to other forms (e.g., XML or CSV) in order to be able to work with it as "data." In doing so, it seemed natural to view the information as coming from three distinct sources, that we think of as a relational database, linked by unique keys in each, as shown in Figure 4.

3.5 Analyzing milestones "data"

Once the milestones data has been re-cast as a database, statistical analysis becomes possible. The simplest case is to look at trends over time. Figure 5 shows a density estimate for the distribution of milestones items from 1500 to the present, keyed to the labels for the periods in history. The bumps, peaks and troughs all seem interpretable: note particularly the steady rise up to $\sim$ 1880, followed by a decline through the "modern dark ages" to $\sim$ 1945, then the steep rise up to the present.

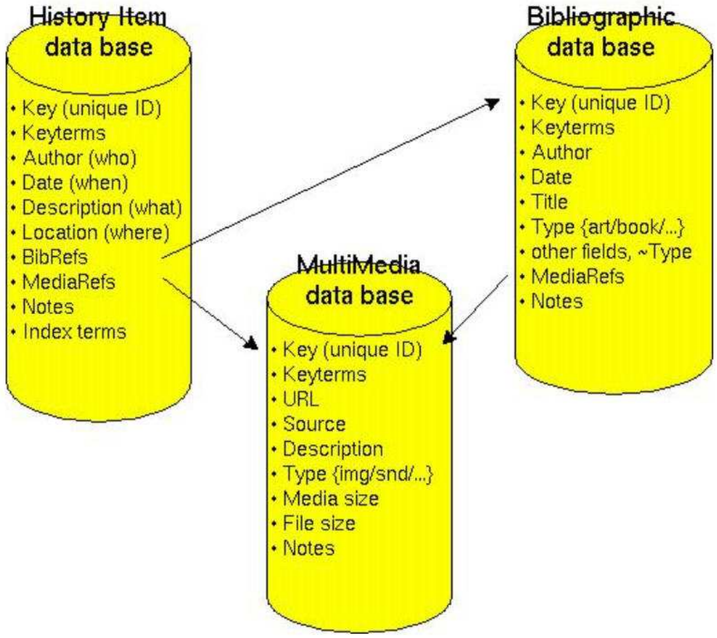

Fig. 4. Milestones data as a relational database composed of history-item, bibliographic, and multimedia databases

If we classify the items by place of development (Europe vs. North America), other interesting trends appear (Figure 6). The decline in Europe following the Golden Age was accompanied by an initial rise in North America, largely due to popularization (e.g., text books) and significant applications of graphical methods, then a steep decline as mathematical statistics held sway.

3.6 What was he thinking?: Understanding through reproduction

Historical graphs were created using available data, methods, technology, and understanding current at the time. We can often come to a better understanding of intellectual, scientific, and graphical questions by attempting a re-analysis from a modern perspective.

Earlier, we showed Playfair's time-series graph (Figure 2) of wages and prices, and noted that Playfair wished to show that workers were better off at the end of the period shown than at any earlier time. Presumably he wished to draw the reader's eye to the narrowing of the gap between the bars for prices and the line graph for wages. Is this what you see?

What this graph shows directly is quite different than Playfair's intension. It appears that wages remained relatively stable, while the price of wheat varied greatly. The inference that wages increased relative to prices is indirect and not visually compelling.

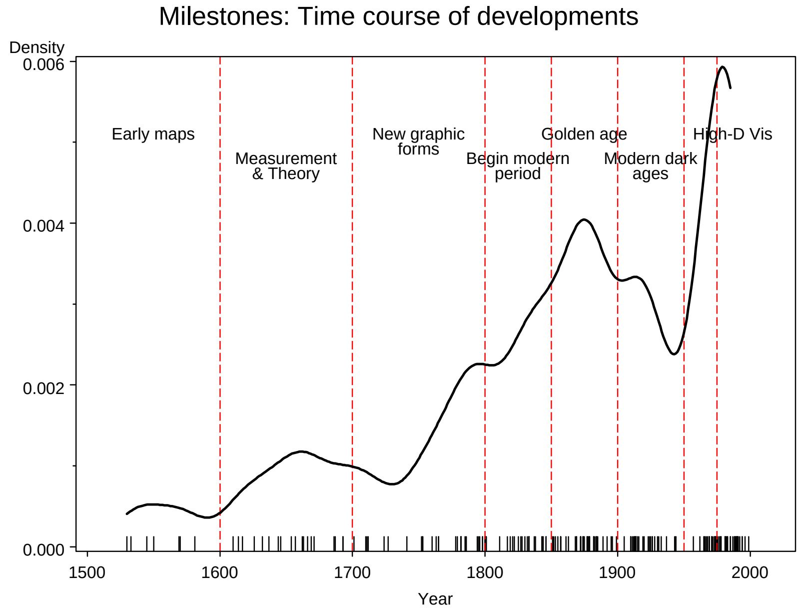

Fig. 5. The distribution of milestone items over time, shown by a rug plot and density estimate.

We cannot resist the temptation to give Playfair a helping hand here—by graphing the ratio of wages to prices (labor cost of wheat), as shown in Figure 7. But this would not have occurred to Playfair, because the idea of relating one time series to another by ratios (index numbers) would not occur for another half-century (Jevons). See Friendly and Denis (2004) for further discussion of Playfair's thinking.

3.7 What kinds of tools are needed?

We have also wondered how other advances in statistics and data visualization could be imported to a historical realm. Among other topics, there has recently been a good deal of work in document analysis and classification that suggests an analog of EDA we might call Exploratory Bibliographic Analysis (EBA).

It turns out that there are several instances of software systems that provide some basic tools for this purpose. An example is RefViz (`http://www.refviz.com`), shown in Figure 8. This software links to common bibliographic software (EndNote, ProCite, Reference Manager, etc.), codes references using key terms from the title and abstract and calculates an index of similarity between pairs of references based on frequencies of co-occurrence. Associations between documents can be shown in a color-coded matrix view, as in Figure 8, or a galaxy view (combining cluster analysis and MDS), and each view offers zoom/unzoom, sorting by several criteria, and querying individual documents or collections.

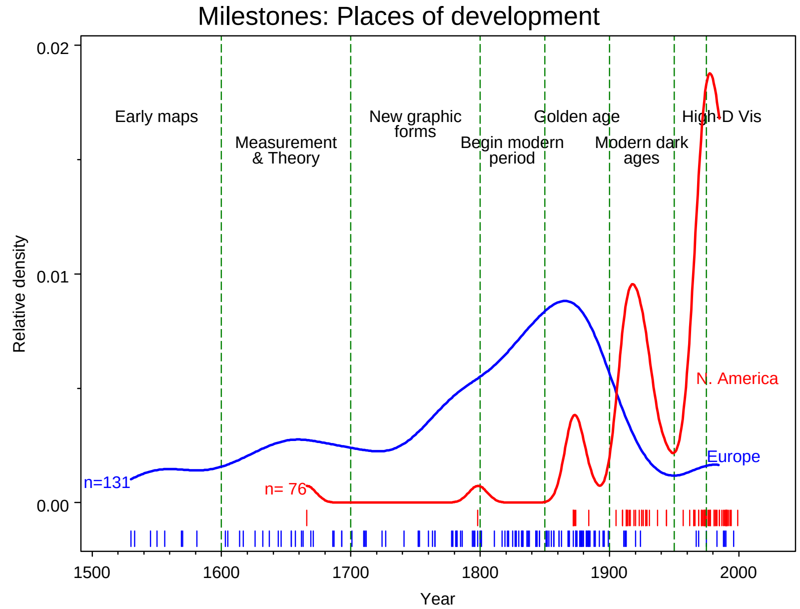

Fig. 6. The distribution of milestone items over time, comparing trends in Europe and North America.

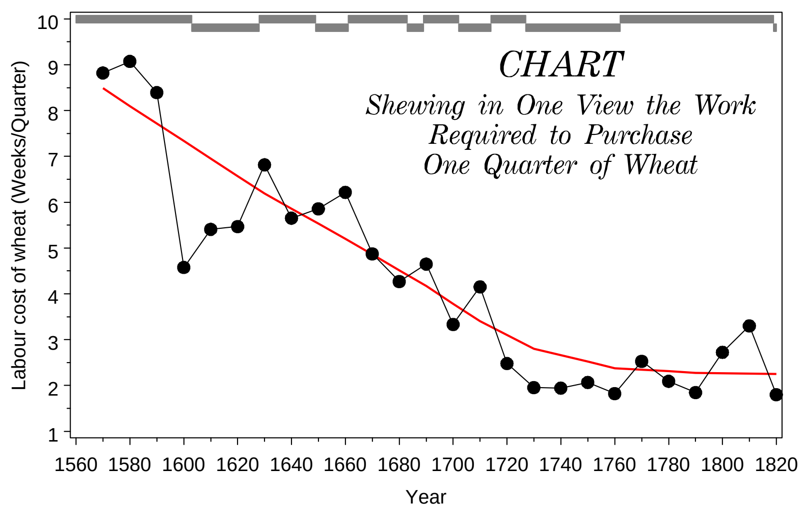

Fig. 7. Redrawn version of Playfair's time series graph showing the ratio of price of wheat to wages, together with a loess smoothed curve.

4 How to visualize a history?

A timeline is obvious, but has severe limitations. We record a history of over 8000 years, but only the last 300-400 have substantial contributions. As well, a linear

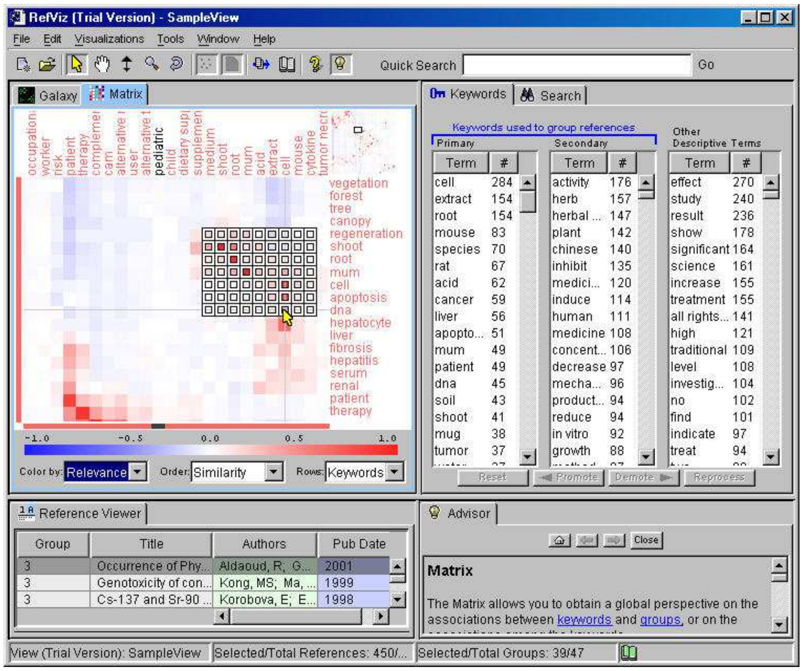

Fig. 8. RefViz similarity matrix view of a bibliographic database. The popup grid is a zoomed display of the region surrounding a selected cell.

representation entails problems of display, resolution and access to detailed information, with little possibility to show either content or context. We explore a few ways to escape these constraints below.

4.1 Lessons from the past

In the milestones collection, we have three examples of attempts to display a history visually. It is of interest that all three used essentially the same format: a horizontal, linear scale for time, with different content or context stacked vertically, as separate horizontal bands.

We illustrate with Joseph Priestley's *Chart of Biography* (Priestley, 1765), showing the lifespans of famous people from 1200 BC to 1750 (Figure 9). Priestley divided people into two groups: 30 "men of learning" and 29 "statesmen," showing each lifespan as a horizontal line. He invented the convention of using dots to indicate uncertainty about exact date of birth or death.

4.2 Lessons from the present

In modern times, a variety of popular publications, mostly in poster form, have attempted to portray graphically various aspects of the history of civilization, geographic regions, or of culture and science.

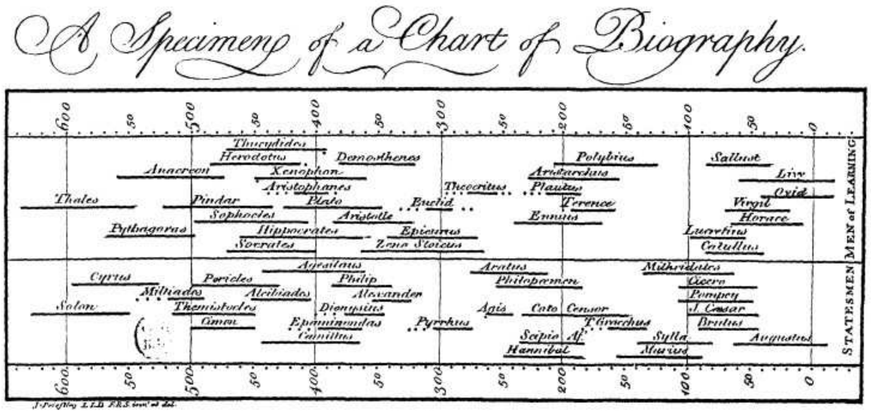

Fig. 9. Priestley's *Chart of Biography*. *Source*: Priestley (1765)

For example, Hammond's *Graphic History of Mankind* (Figure 10) shows the emergence of new cultures and the rise and fall of various empires, nations and ethnic groups from the late Stone Age to the present in a vertical format. It uses a varying-resolution time scale, quite coarse in early history, getting progressively finer up to recent times. It portrays these using flow lines of different colors, whose width indicates the influence of that culture, and with shading or stripes to show conquest or outside influence.

4.3 Lessons from the web

A large component of the milestones collection is the catalog of graphic images and portraits associated with the milestones items. At present, they are stored as image files of fixed resolution and size, and presented as hyper-links in the public versions. How can we do better, to make this material more easily accessible?

There are now a number of comprehensive image libraries available on the web that provide facilities to search for images by various criteria and in some cases to view these at varying resolutions. Among these, David Rumsey's Map Collection (`http://www.davidrumsey.com`) is notable. It provides access to a collection of over 8800 historical maps (mostly 18^{th}-19^{th} century, of North/South America, with some European content) online, extensively indexed so they may be searched by author, category, country or region, and a large number of other data fields. The maps are stored using Mr. Sid technology (`http://www.lizardtech.com`), which means that they can be zoomed and panned in real time. Rumsey provides several different browsers, including a highly interactive Java client.

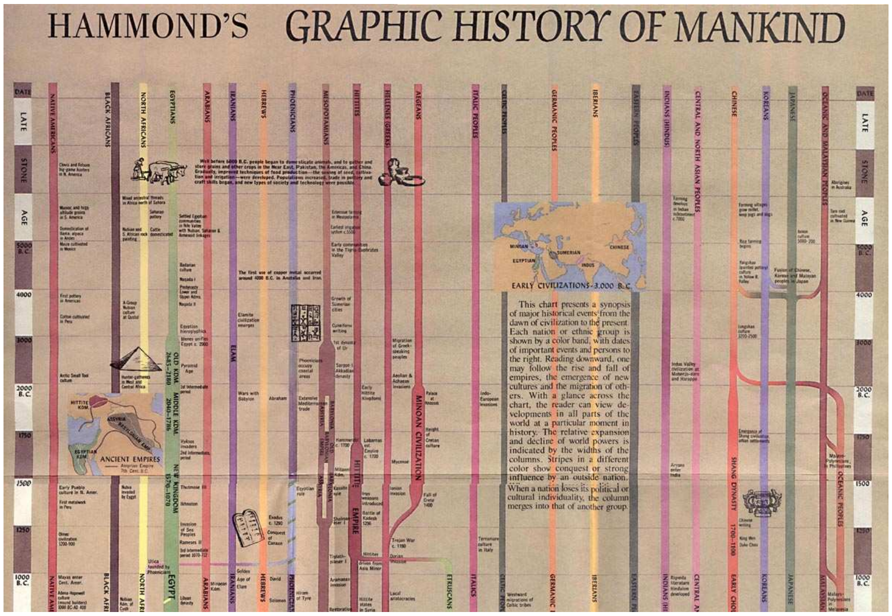

Fig. 10. Hammond's *Graphic History of Mankind* (first of 5 panels)

4.4 Lessons from the data visualization

Modern data visualization also provides a number of different ideas and approaches to multivariate complexity, time and space we may adapt (in a self-referential way) to the history of data visualization itself.

Interactive viewers provide one simple solution to the trade-off between detail and scope of a data view through zoom and unzoom, but in the most basic implementation, any given view is a linear scaling of the section of the timeline that will fit within the given window.

We can do better by varying resolution *continuously* as a non-linear, decreasing function of distance from the viewer's point of focus. For example, Figure 11 shows a fisheye view (Furnas, 1986) of central Washington, D.C., using a hyperbolic scale, so that resolution is greatest at the center and decreases as 1/distance. The map is dynamic, so that moving the cursor changes the focal point of highest resolution, This has the property that it allows the viewer to see the context surrounding the point of focus, yet navigate smoothly throughout the entire space. Similar ideas have been applied to tables in the Table Lens (`http://www.tablelens.com`) and hierarchies (Lamping *et al.*, 1995) such as web sites and file systems, and can easily be used for a 1D timeline.

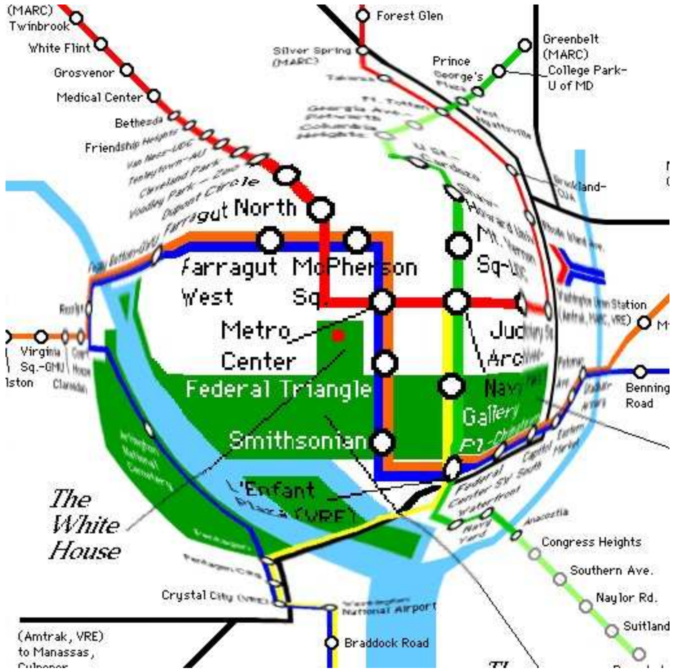

Fig. 11. Fisheye view of central Washington, D.C., illustrating a hyperbolic view

Acknowledgments

This paper is based on joint work with Dan Denis, now at the University of Montana. It was supported by Grant 8150 from the National Science and Research Council of Canada and was prepared with the assistance of *Les Chevaliers des Album de Statistique Graphique*, particularly Antoine de Falguerolles and Rüdiger Ostermann, to whom we are grateful. We also thank several reviewers for helping to make this presentation more cogent.

Bibliography

BECKER, R. A. (1994): A brief history of S. In P. Dirschedl and R. Ostermann, eds., *Computational Statistics*, (pp. 81–110). Heidleberg: Physica Verlag.

BENIGER, J. R. AND ROBYN, D. L. (1978): Quantitative graphics in statistics: A brief history. *The American Statistician*, 32, 1–11.

BERTIN, J. (1967): *Sémiologie Graphique: Les diagrammes, les réseaux, les cartes*. Paris: Gauthier-Villars.

BERTIN, J. (1983): *Semiology of Graphics*. Madison, WI: University of Wisconsin Press. (trans. W. Berg).

FAREBROTHER, R. W. (1999): *Fitting Linear Relationships: A History of the Calculus of Observations 1750–1900*. New York: Springer.

FRIENDLY, M. (1994): Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, 89, 190–200.

FRIENDLY, M. (2000): Re-Visions of Minard. *Statistical Computing & Statistical Graphics Newsletter*, 11(1), 1, 13–19.

FRIENDLY, M. AND DENIS, D. (2000): The roots and branches of statistical graphics. *Journal de la Société Française de Statistique*, 141(4), 51–60. (published in 2001).

FRIENDLY, M. AND DENIS, D. (2004): The early origins and development of the scatterplot. *Journal of the History of the Behavioral Sciences*. (In press, accepted 7/09/04).

FRIENDLY, M. AND KWAN, E. (2003): Effect ordering for data displays. *Computational Statistics and Data Analysis*, 43(4), 509–539.

FRIIS, H. R. (1974): Statistical cartography in the United States prior to 1870 and the role of Joseph C. G. Kennedy and the U.S. Census Office. *American Cartographer*, 1, 131–157.

FUNKHOUSER, H. G. (1936): A note on a tenth century graph. *Osiris*, 1, 260–262.

FUNKHOUSER, H. G. (1937): Historical development of the graphical representation of statistical data. *Osiris*, 3(1), 269–405. Reprinted Brugge, Belgium: St. Catherine Press, 1937.

FURNAS, G. W. (1986): Generalized fisheye views. In *Proceedings of the ACM CHI '86 Conference on Human Factors in Computing Systems*, (pp. 16–23). Boston, MA: ACM.

HALD, A. (1990): *A History of Probability and Statistics and their Application before 1750*. New York: John Wiley and Sons.

HANKINS, T. L. (1999): Blood, dirt, and nomograms: A particular history of graphs. *Isis*, 90, 50–80.

HEISER, W. J. (2000): Early roots of statistical modelling. In J. Blasius, J. Hox, E. de Leeuw, and P. Schmidt, eds., *Social Science Methodology in the New Millenium: Proceedings of the Fifth International Conference on Logic and Methodology*. Amsterdam: TT-Publikaties.

HOFF, H. E. AND GEDDES, L. A. (1959): Graphic recording before Carl Ludwig: An historical summary. *Archives Internationales d'Histoire des Sciences*, 12, 3–25.

HOFF, H. E. AND GEDDES, L. A. (1962): The beginnings of graphic recording. *Isis*, 53, 287–324. Pt. 3.

KRUSKAL, W. (1977): Visions of maps and graphs. In *Proceedings of the International Symposium on Computer- Assisted Cartography, Auto-Carto II*, (pp. 27–36). 1975.

LAMPING, J., RAO, R., AND PIROLLI, P. (1995): A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, (pp. 401–408). ACM Press.

MERTON, R. K. (1973): *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: University of Chicago Press.

MINARD, C. J. (1844): Tableaux figuratifs de la circulation de quelques chemins de fer. lith. (n.s.). ENPC: 5860/C351, 5299/C307.

PALSKY, G. (1996): *Des Chiffres et des Cartes: Naissance et développement de la Cartographie Quantitative Français au XIX^e siècle*. Paris: CTHS.

PEARSON, E. S., ed. (1978): *The History of Statistics in the 17th and 18th Centuries Against the Changing Background of Intellectual, Scientific and Religeous Thought*. London: Griffin & Co. Ltd. Lectures by Karl Pearson given at University College London during the academic sessions 1921–1933.

PLAYFAIR, W. (1786): *Commercial and Political Atlas: Representing, by Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure, and Debts of England, during the Whole of the Eighteenth Century*. London: Corry. 3rd edition, London: Stockdale, 1801; French edition, Tableaux d'arithmétique linéaire, du commerce, des finances, et de la dette nationale de l'Angleterre (Paris: Chez Barrois l'Aîné, 1789).

PLAYFAIR, W. (1801): *Statistical Breviary; Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe*. London: Wallis.

PLAYFAIR, W. (1821): Letter on our agricultural distresses, their causes and remedies; accompanied with tables and copperplate charts shewing and comparing the prices of wheat, bread and labour, from 1565 to 1821.

PORTER, T. M. (1986): *The Rise of Statistical Thinking 1820–1900*. Princeton, NJ: Princeton University Press.

PRIESTLEY, J. (1765): A chart of biography. London.

RIDDELL, R. C. (1980): Parameter disposition in pre-Newtonain planetary theories. *Archives Hist. Exact Sci.*, 23, 87–157.

ROBINSON, A. H. (1982): *Early Thematic Mapping in the History of Cartography*. Chicago: University of Chicago Press.

ROYSTON, E. (1970): Studies in the history of probability and statistics, III. a note on the history of the graphical presentation of data. *Biometrika*, (pp. 241–247). 43, Pts. 3 and 4 (December 1956); reprinted In *Studies in the History Of Statistics and Probability Theory*, eds. E. S. Pearson and M. G. Kendall, London: Griffin.

STIGLER, S. M. (1980): Stigler's law of eponomy. *Transactions of the New York Academy of Sciences*, 39, 147–157.

STIGLER, S. M. (1986): *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Harvard University Press.

TILLING, L. (1975): Early experimental graphs. *British Journal for the History of Science*, 8, 193–213.

TUFTE, E. R. (1983): *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

TUFTE, E. R. (1990): *Envisioning Information*. Cheshire, CT: Graphics Press.

TUFTE, E. R. (1997): *Visual Explanations*. Cheshire, CT: Graphics Press.

WALLIS, H. M. AND ROBINSON, A. H. (1987): *Cartographical Innovations: An International Handbook of Mapping Terms to 1900*. Tring, Herts: Map Collector Publications.

WHEELER, J. A. (1982): Bohr, Einstein, and the strange lesson of the quantum. In R. Q. Elvee, ed., *Mind in Nature*. San Francisco: Harper and Row.