History Of Data Visualization

Idea In Short

Data visualization has come a long way. From simple cave drawings depicting the success of a hunt to the intricate dashboards, the history of data visualization is full of incredible stories marked by major events2.

Michael Friendly from York University published a paper on the milestones in the history of Data Visualization, Milestones in the History of Data Visualization: A Case Study in Statistical Historiography3. The paper provides an excellent breakdown of major points that have lead us to modern day data visualization. Prior to the 17th century, data visualization existed mainly in the realm of maps, displaying land markers, cities, roads, and resources. As the demand grew for more accurate mapping and physical measurement, better visualizations were needed4.

Pre-history

The first visualizations5 may have taken the form of drawings in sand or scratched on rock and it is possible that the famous Palaeolithic cave paintings in Lascaux, southern France, may have functioned as both hunting guides and directions to the spirit world. The ancient Babylonians, Egyptians, Greeks and Chinese all developed sophisticated ways of representing information visually to plot the movements of the stars, produce maps to aid navigation, and develop plans for crop planting and city development.

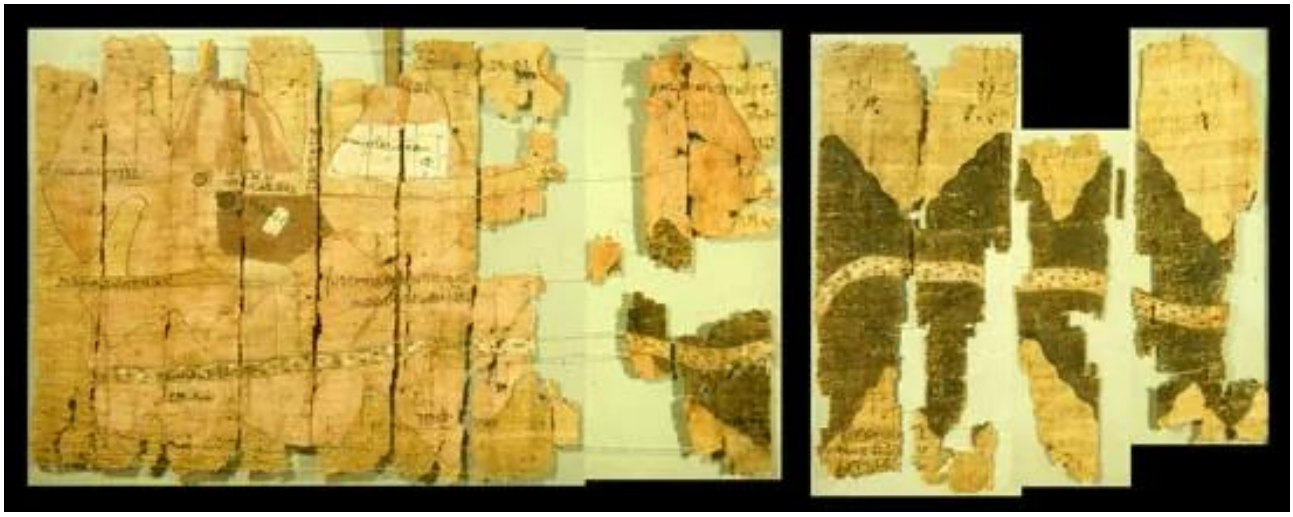


Babylonian Maps

Many of these early visualizations would have been drawn on clay such as this very early Babylonian world map (600 BC). Later visualizations would be rendered onto papyrus, which had the benefit of enabling information to be more easily shared and annotated.

Turin Papyrus Map (Egypt, 1150 BC)

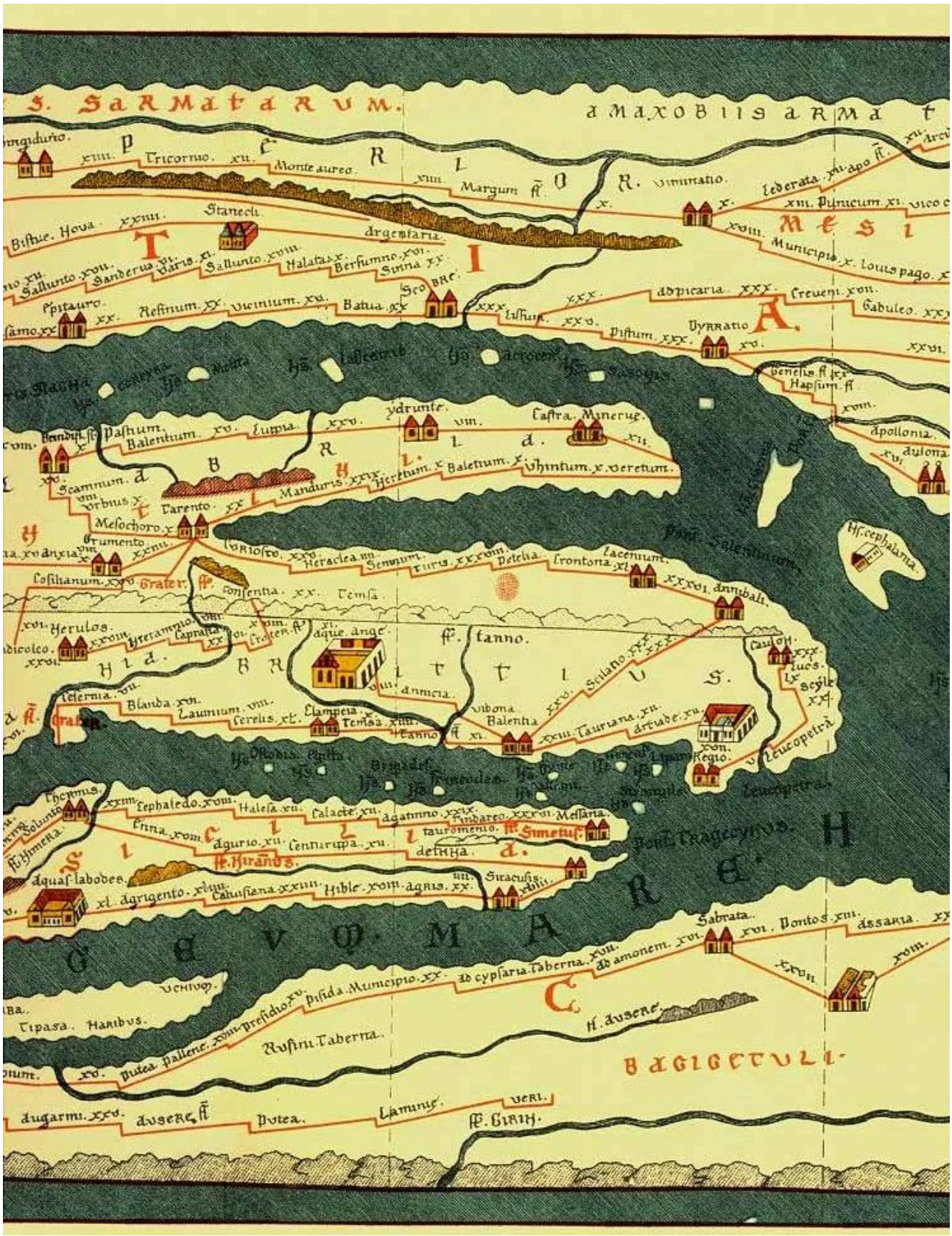
Found in Thebes in the 19th century, this sophisticated Egyptian papyrus map displays a fascinating range of information including geographical and mineral mining data. Recent studies have shown that its colour-coded geological information is remarkably accurate.



Turin Papyrus

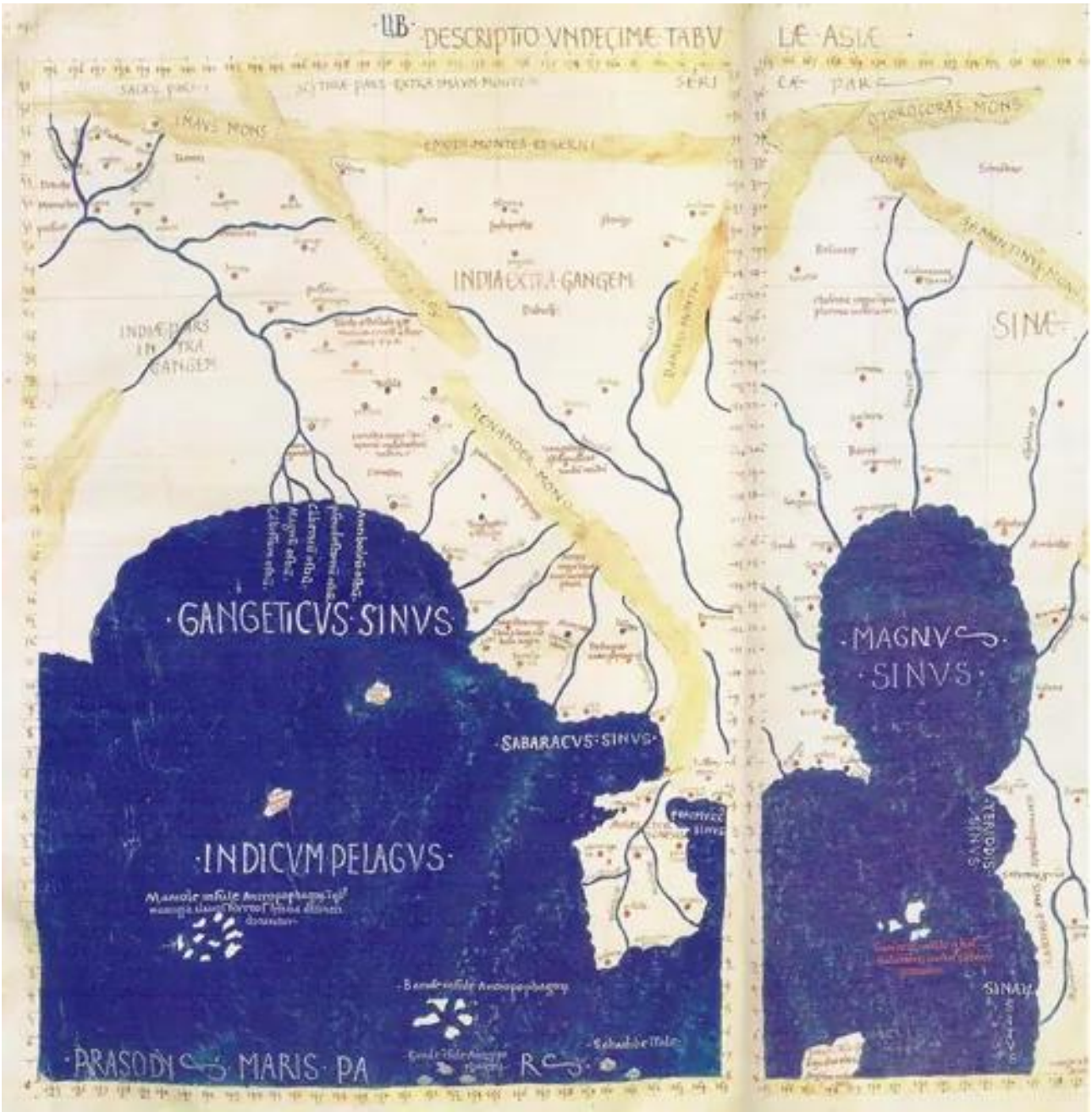
Roman Maps, the Tabula Peutinger (366-335 BC)

The Romans were renowned mapmakers, as they needed to efficiently plan for the movement of their armies and trade throughout the empire. Although only shown here in detail the Peutinger map is a 7 metre long reproduction made by a monk in the thirteenth century. It displays the road system of the empire starting in Britain in the west and finishing in India in the east. Routes are marked as lines and destinations displayed as icons. In this sense it similar to a schematic diagram like the London Underground map designed by Henry Beck in 1931.



Ptolemy's World Map (2nd century AD)

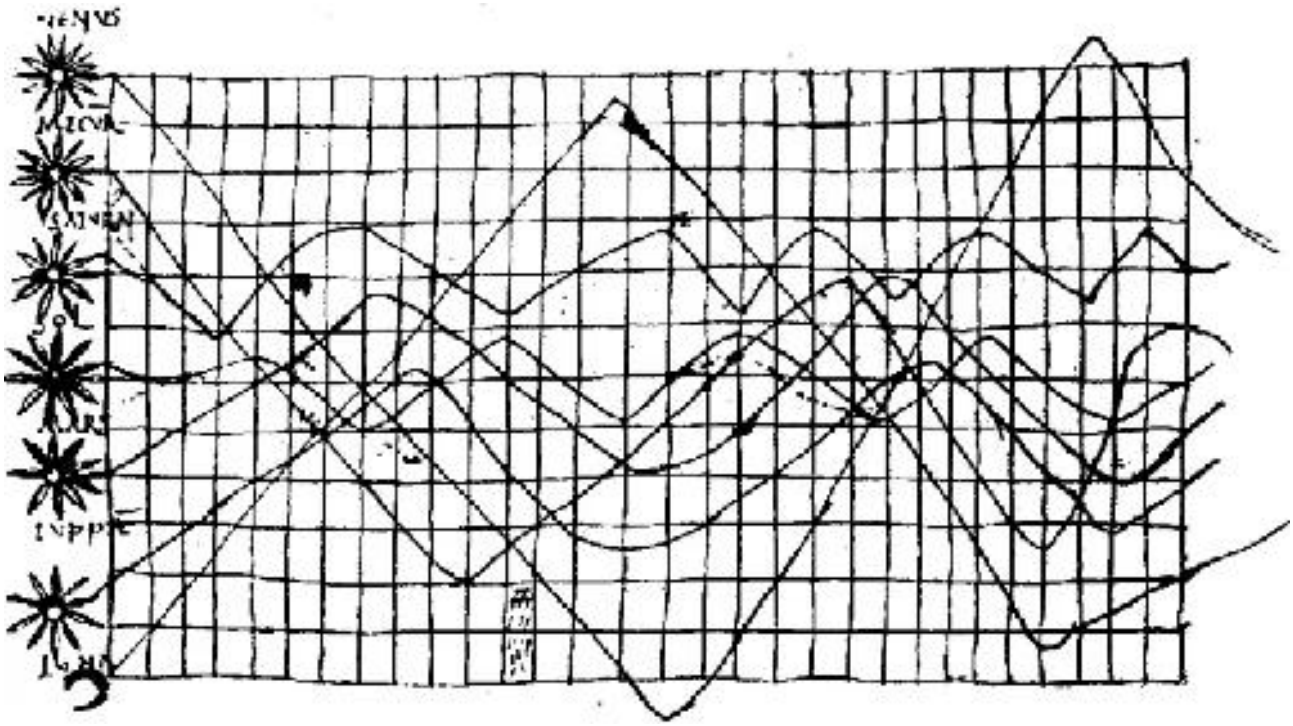
Later maps such as Ptolemy's world map would revolutionize map-making through the use of latitude and longitude markings to develop a global coordinate system. By applying scientific method, maths and astronomy to specify location, these images are forerunners of modern maps.



Ptolemy's World Map

Positions of the Sun, Moon, and Planets Throughout the Year (Europe, 950 AD)

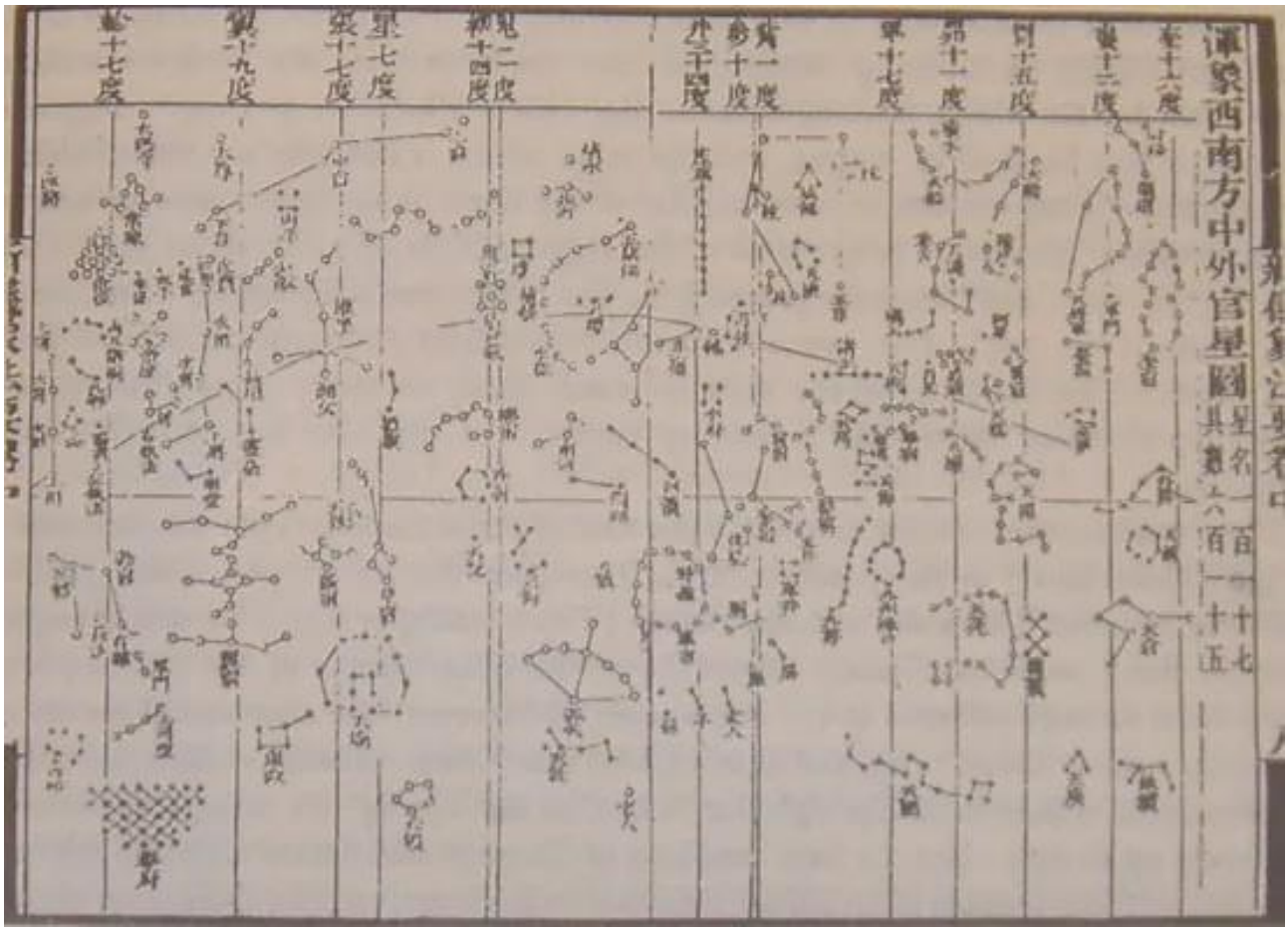
This fascinating diagram is reminiscent of a modern information visualization as it uses a grid system to combine both the time and location of planetary bodies 800 years before the first true time plotted graphics. While its exact meaning is now difficult to discern, the right-hand side clearly indicates the position and trajectory of the planets and stars, while the horizontal grid refers to periods of time.



Positions Of The Sun

Su Song's Celestial Atlas (China, 1092 AD)

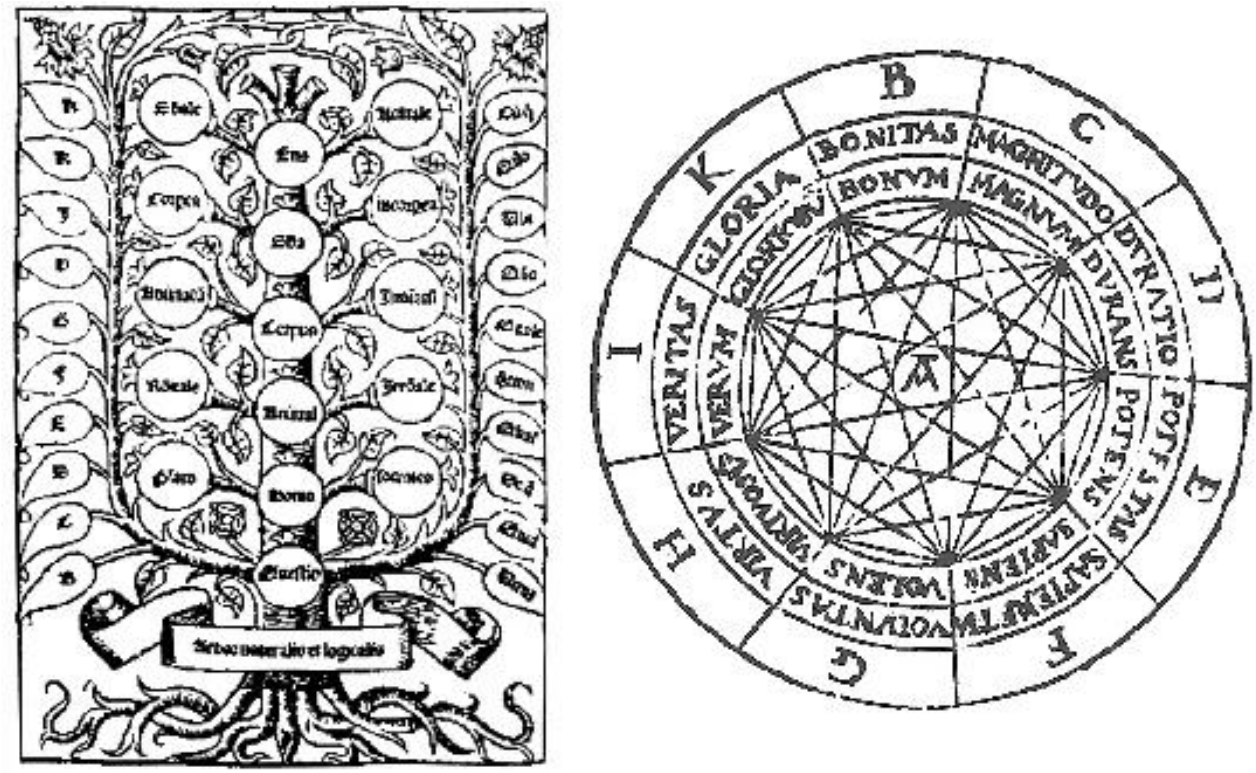
Su Song was a Chinese scientist active during the Song Dynasty. This star chart uses a sophisticated projection of celestial bodies that employs techniques not introduced to Europe until the 16th century. While this chart was produced after the European chart described above, the ancient Chinese had been producing highly sophisticated visualizations of the stars for nearly a millennia.



Su Song Star Map

Ramon Llull, Diagrams of Relationships Between Knowledge

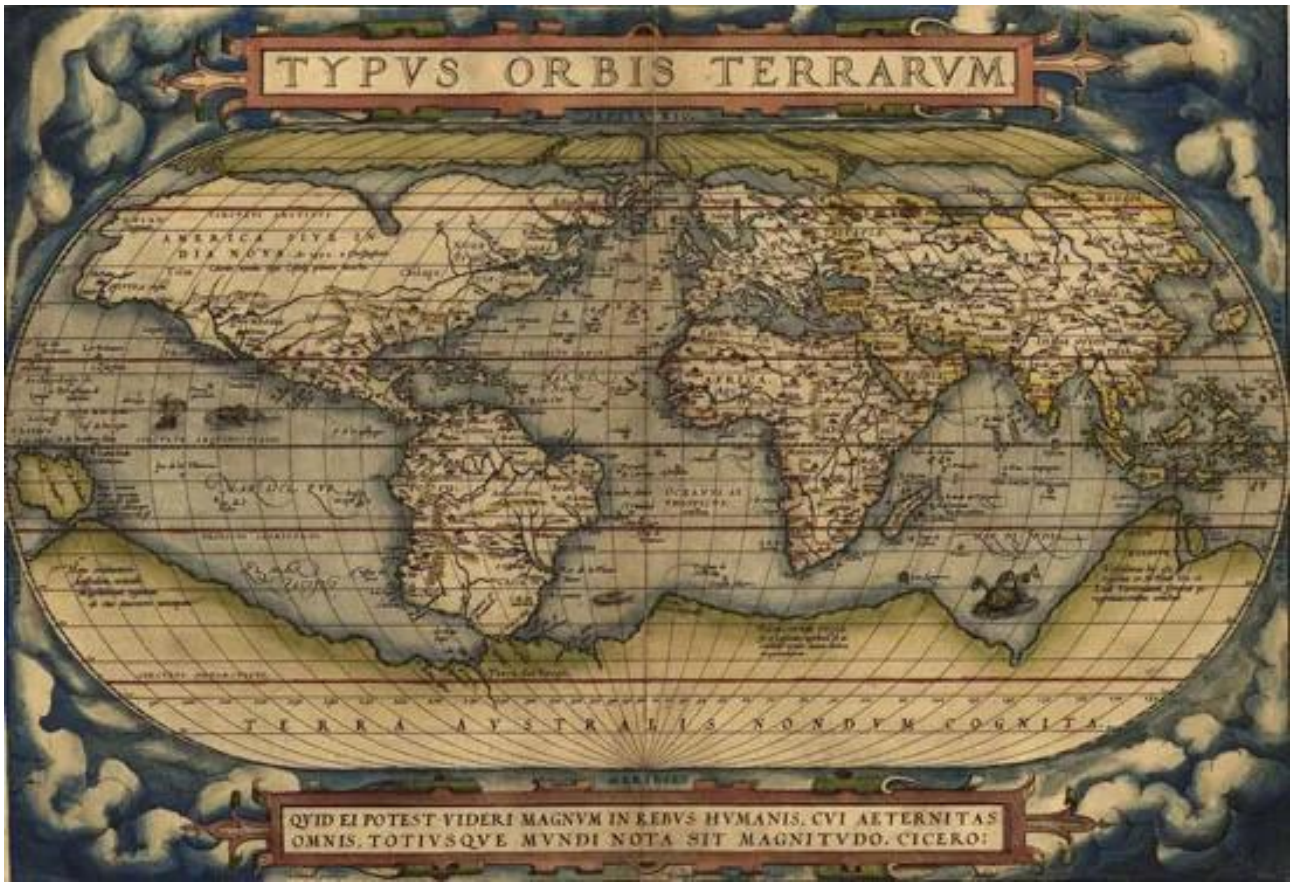
Llull was a Spanish philosopher who lived between the 13th and 14th centuries. He believed that there were basic concepts common to all experience, and that by arranging these ideas diagrammatically truths and insights into the human condition could be generated. The wheel diagram shows a series of rotating disks surrounded by terms. Moving the discs allows a realignment of the terms and new combinations of "truth" to be produced. The tree-like structure similarly shows relationships between different orders of knowledge. The use of images to combine and relate information in different ways to generate insights is a mainstay of contemporary visualisation practices and is often described as "discovery".



Diagrams of Relationships Between Knowledge

Abraham Ortelius, the First Modern Atlas (1570, Belgium)

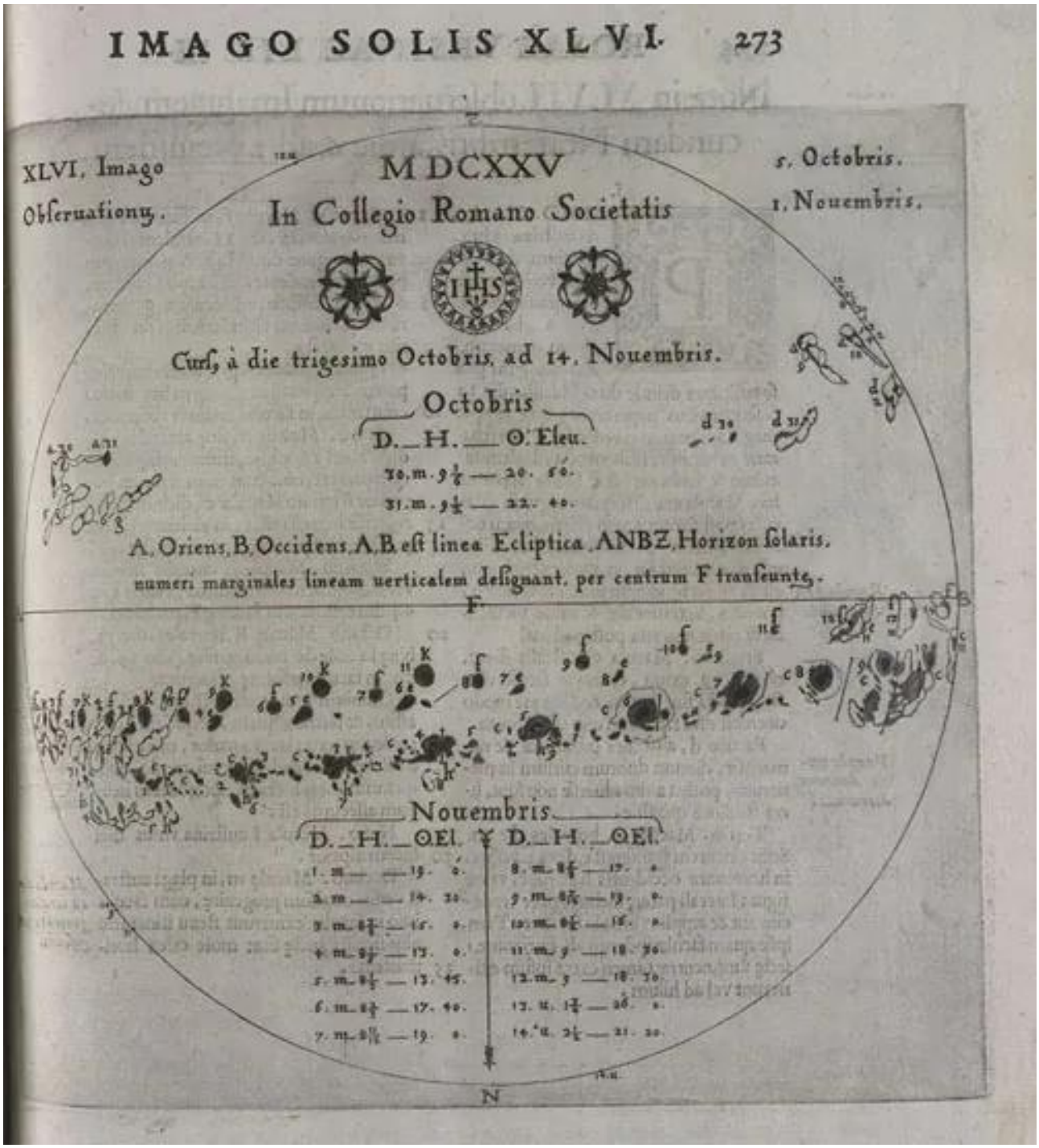
Produced by Ortelius in Antwerp in 1570, the Theatrum Orbis Terrarum (Theatre of the World) contained 53 maps with a comprehensive descriptive text and supplementary material. It's importance lies in the fact that it was the first attempt to gather all known information and maps of the word in one printed volume.



First Modern Atlas

Christoph Scheiner, Images of Sunspots

Scheiner was a Jesuit priest and mathematician working in Germany and Italy in the 17th century. Using state of the art telescopes he studied the markings on the Sun [sunspots] in order to reconcile Church teachings that the Cosmos contained no imperfections, with the visual evidence that it did. The drawings reproduced here demonstrate the changes in sunspots over time and identify different types.

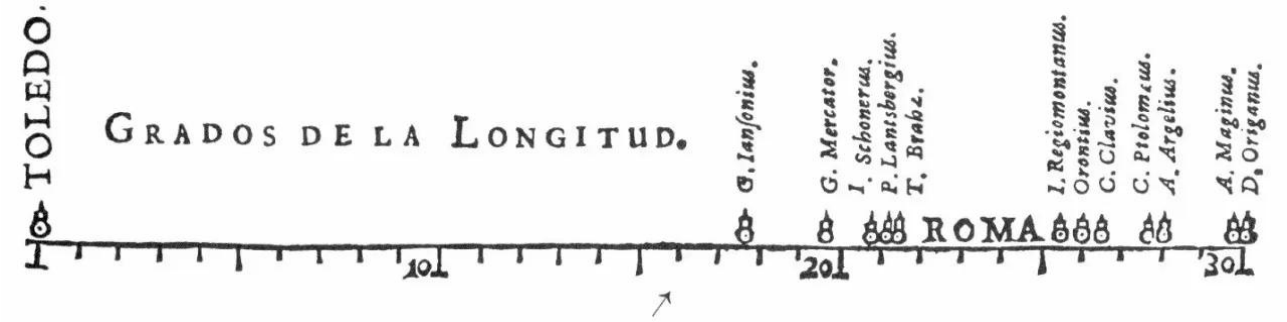


Scheiner Sunspots

Michael Florent van Langren, Visual representation of statistical data

In 1644, Michael Florent van Langren, a Flemmish astronomer, is believed to have provided the first visual representation of statistical data. The one dimensional line graph below shows the twelve known estimates at the time of the difference in longitude between Toledo

and Rome as well as the name of each astronomer who provided the estimate. While Van Langren could have provided this information in a table, it is the use of the graph that really visually displays the wide variations in estimates.



Florent Van Langren

Edmond Halley, Contour Maps

Edmond Halley was an English astronomer, geophysicist, mathematician, meteorologist, and physicist who is best known for computing the orbit of Halley's Comet. According to the BBC, Halley developed the use of contour lines on maps to connect and describe areas that display differences in atmospheric conditions from place to place. These lines are now commonly used to describe meteorological variation common to us from weather reports. The 18th century saw the beginning of thematic mapping. Attempts at thematic mapping of geologic, economic, and medical data were made near the end of the century. Abstract graphs of functions, measurement error, and collection of empirical data were introduced at this time. Many statistical chart types, including histograms, time series plots, contour plots, scatterplots, and others were invented during this period.

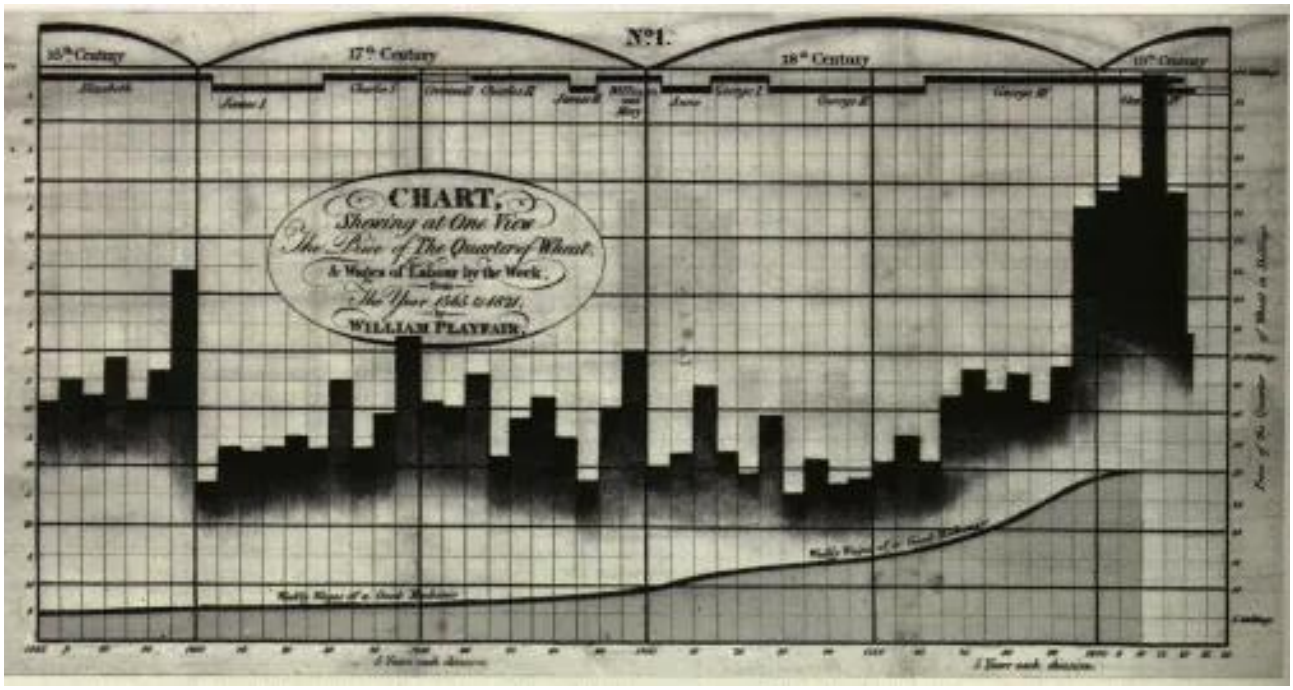


Edmond Halley, Contour Maps

William Playfair, Statistical Graphics

This period also gave us William Playfair, widely considered the the father of statistical graphics, including most popular graphs we use today (line, bar, circle, and pie charts). Playfair was a Scottish engineer and political economist who published The Commercial and

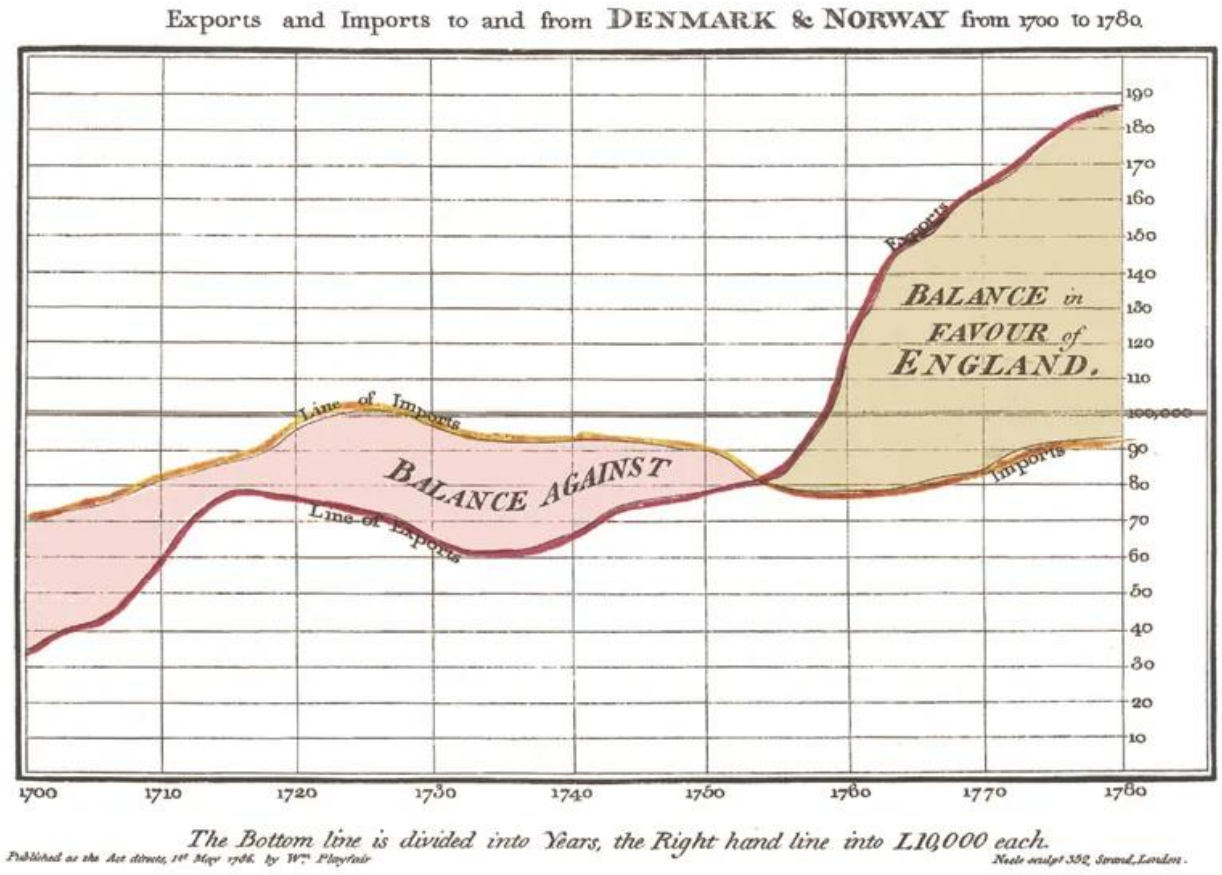
Political Atlas in 1786. A graph by Playfair (1821), shown below, shows the price of wheat, weekly wages, and reigning monarch over a two hundred fifty year span from 1565 to 1820.



William Playfair's Statistical Graphics

This book featured a variety of graphs, including the following image.

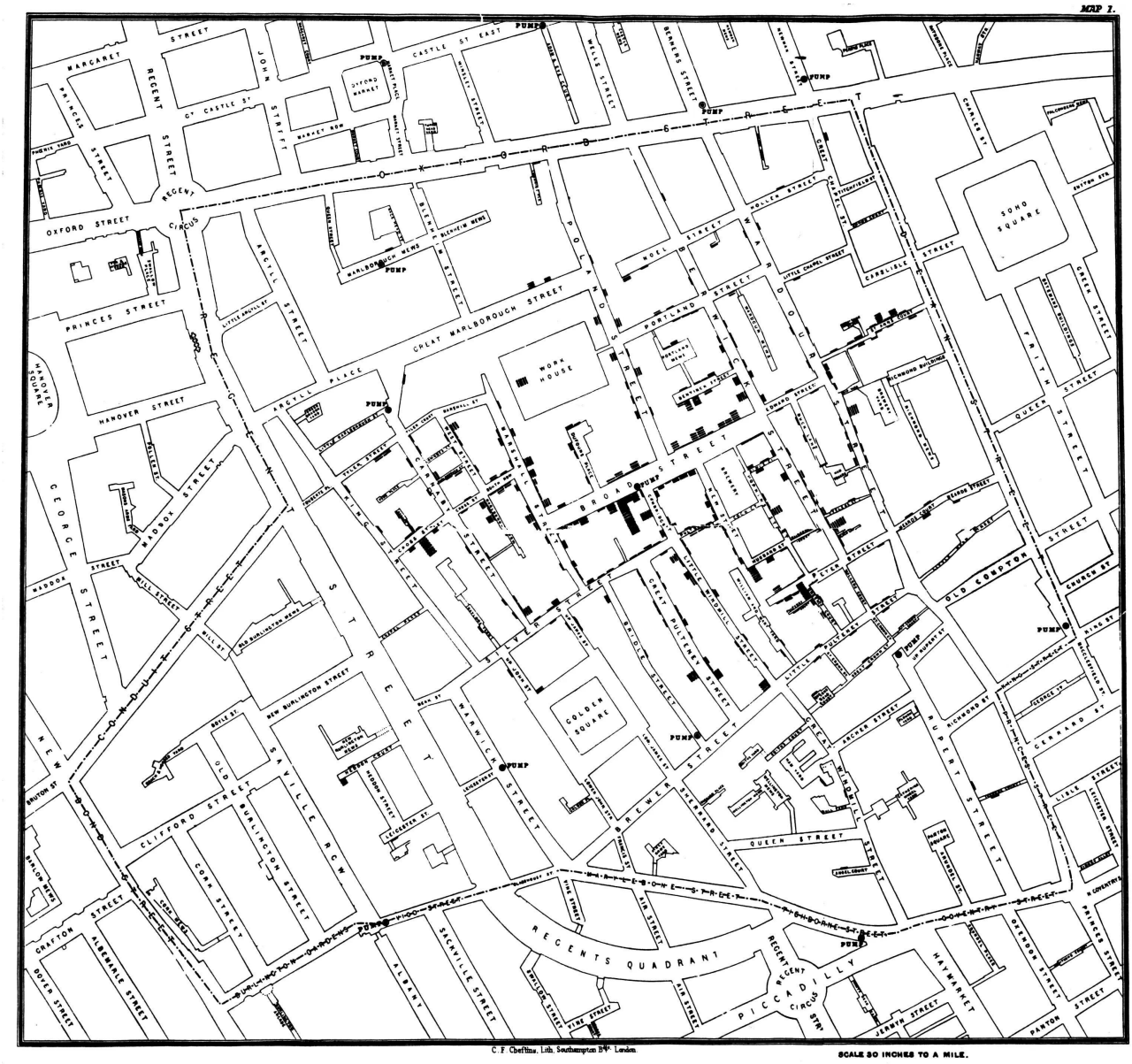
In this famous example, he compares exports from England with imports into England from Denmark and Norway from 1700 to 1780.



Playfair Imports Exports Visualization

John Snow, Map of Cholera outbreaks

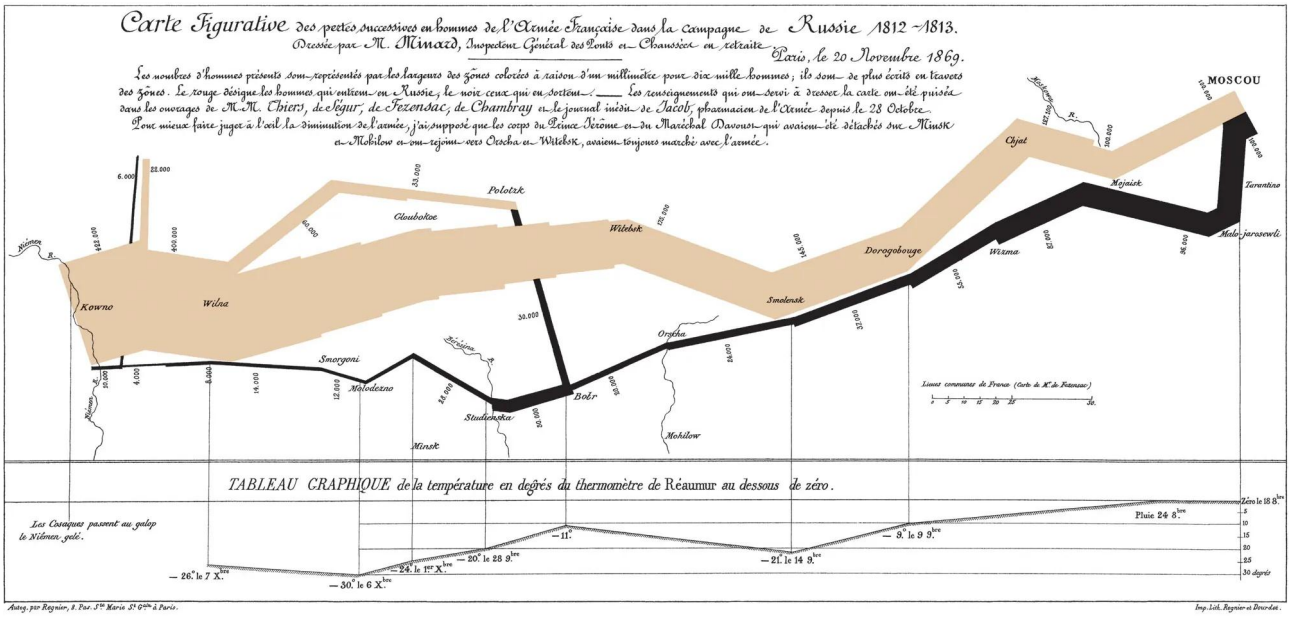
The latter half of the 19th century is what Friendly calls the Golden Age of statistical graphics. Two famous examples of data visualization from that era include John Snow's map of Cholera outbreaks in the London epidemic of 1854.



John Snow - Cholera Outbreaks Visualization

Charles Minard, Russian Campaign

Charles Minard's 1869 chart showing the number of men in Napoleon's 1812 infamous Russian campaign army, with army location indicated by the X-axis, and extreme cold temperatures indicated at points when frostbite took a fatal toll.

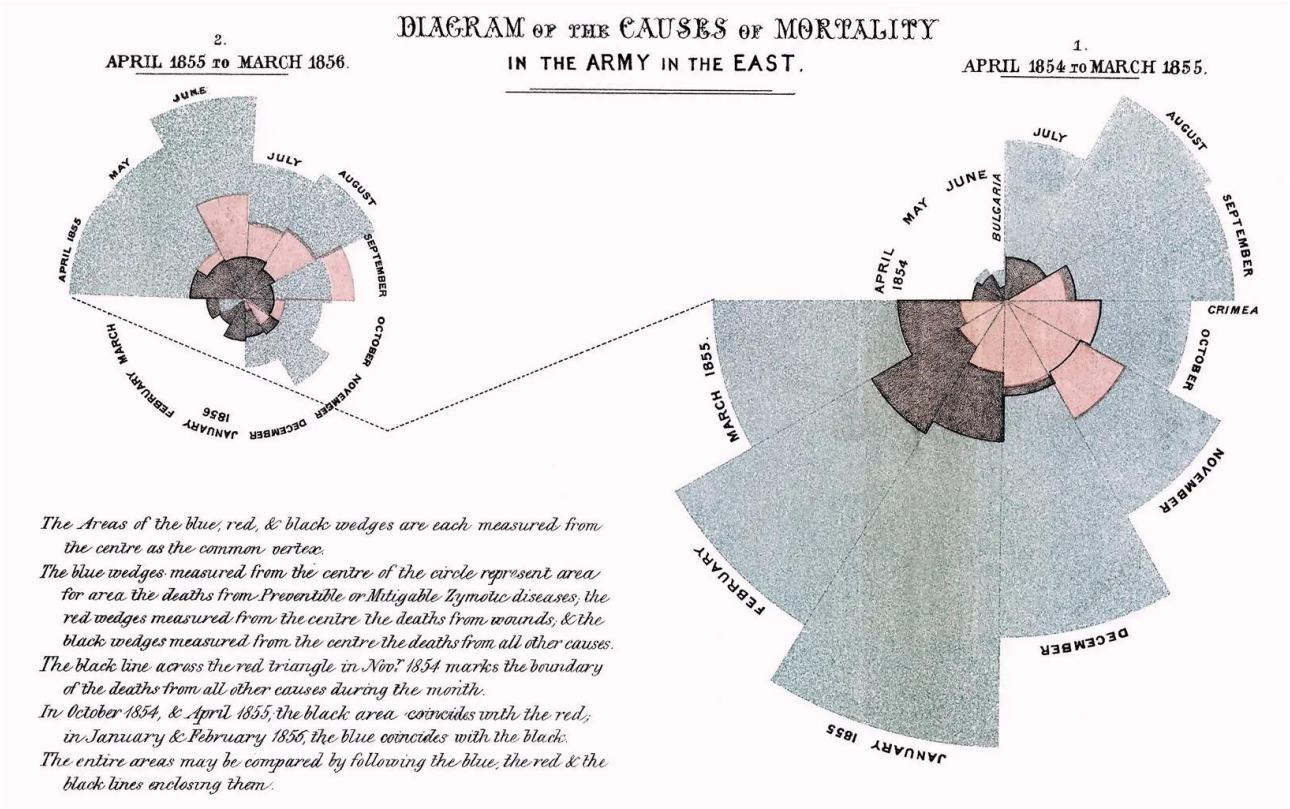


Charles Minard's Russian Campaign Visualization

Florence Nightingale, Rose Chart

This time also provided us with a new visualization, the Rose Chart, created by Florence Nightingale. A number of factors contributed to this "Golden Age" of statistical graphing:

1. The industrial revolution, which created the modern business
2. Official government statistical offices, to support an increasingly aware and global populace
3. A growing recognition for the importance of numerical data in social planning, medicine, military, industrialization, commerce, and transportation

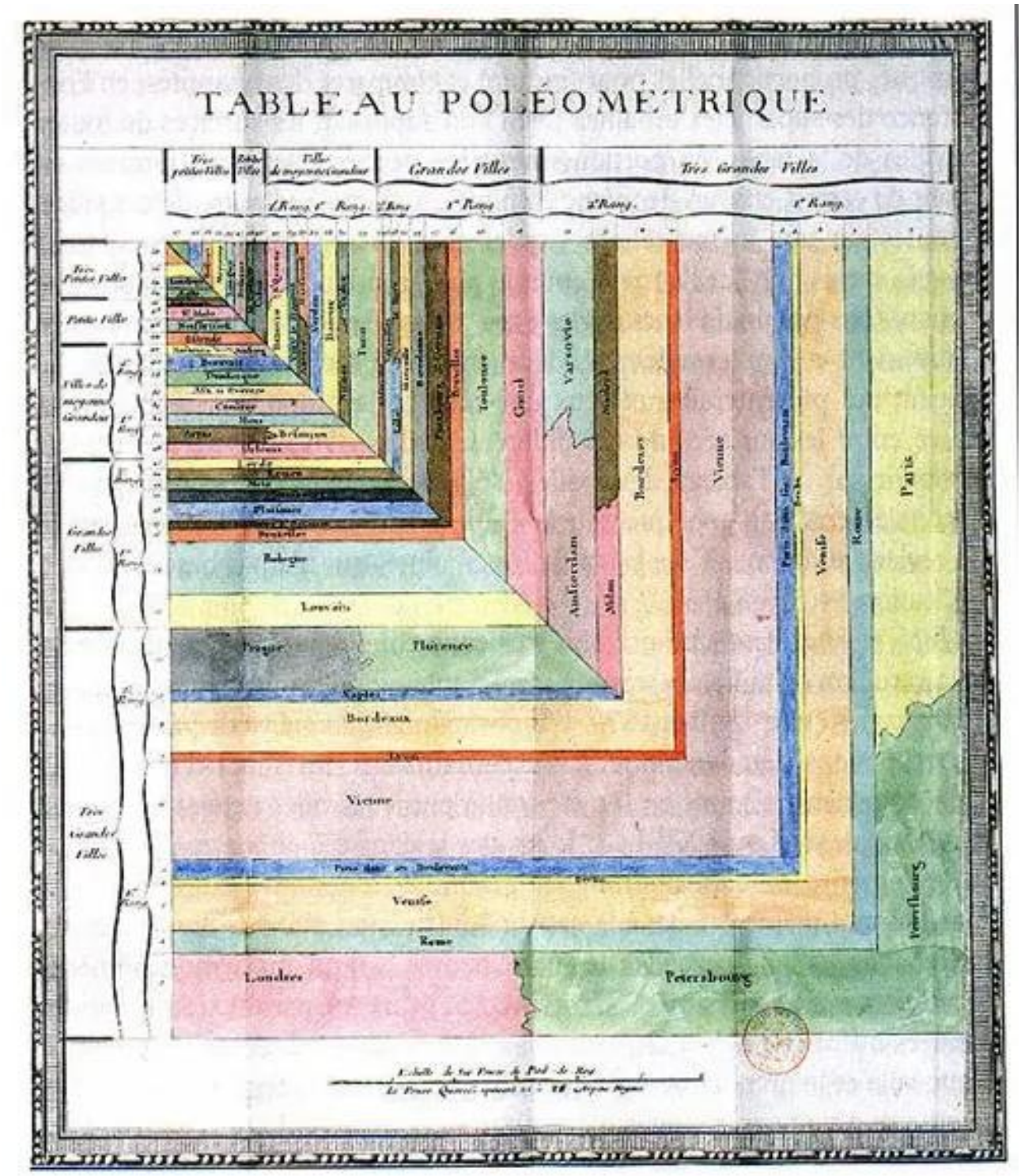


Florence Nightingale's Rose Chart

Statistical Theory also provided the means to make sense of large datasets.

Charles de Fourcroy, Tableau Poléometrique

Charles de Fourcroy was a French mathematician and scholar. He produced a visual analysis of the work of French civil engineers and a comparison of the demographics of European cities. In 1782 he published Tableau Poléometrique, a treatise on engineering and civil construction. His use of geometric shapes predates the modern treemap, which is widely used today to display hierarchical data.

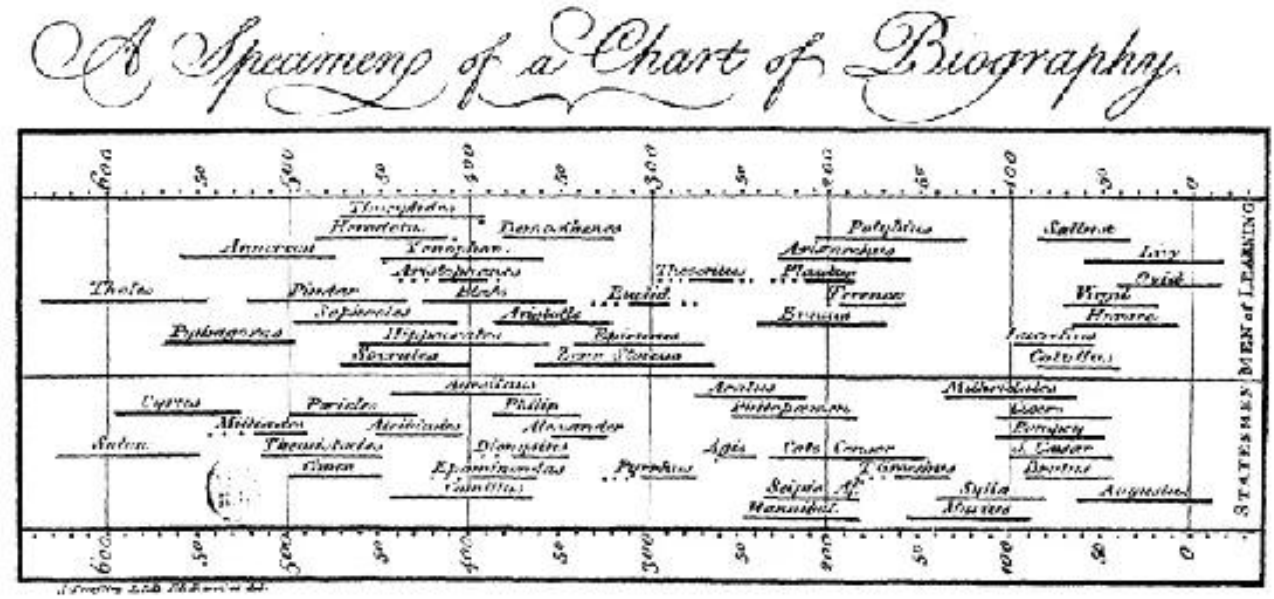


Charles de Fourcroy's Tableau Poléometrique

Priestley, Life span of famous persons

An English chemist and theologian Priestley produced a biographical chart comparing the life spans of the famous stretching back to pre-history. The chart maps names to bars and

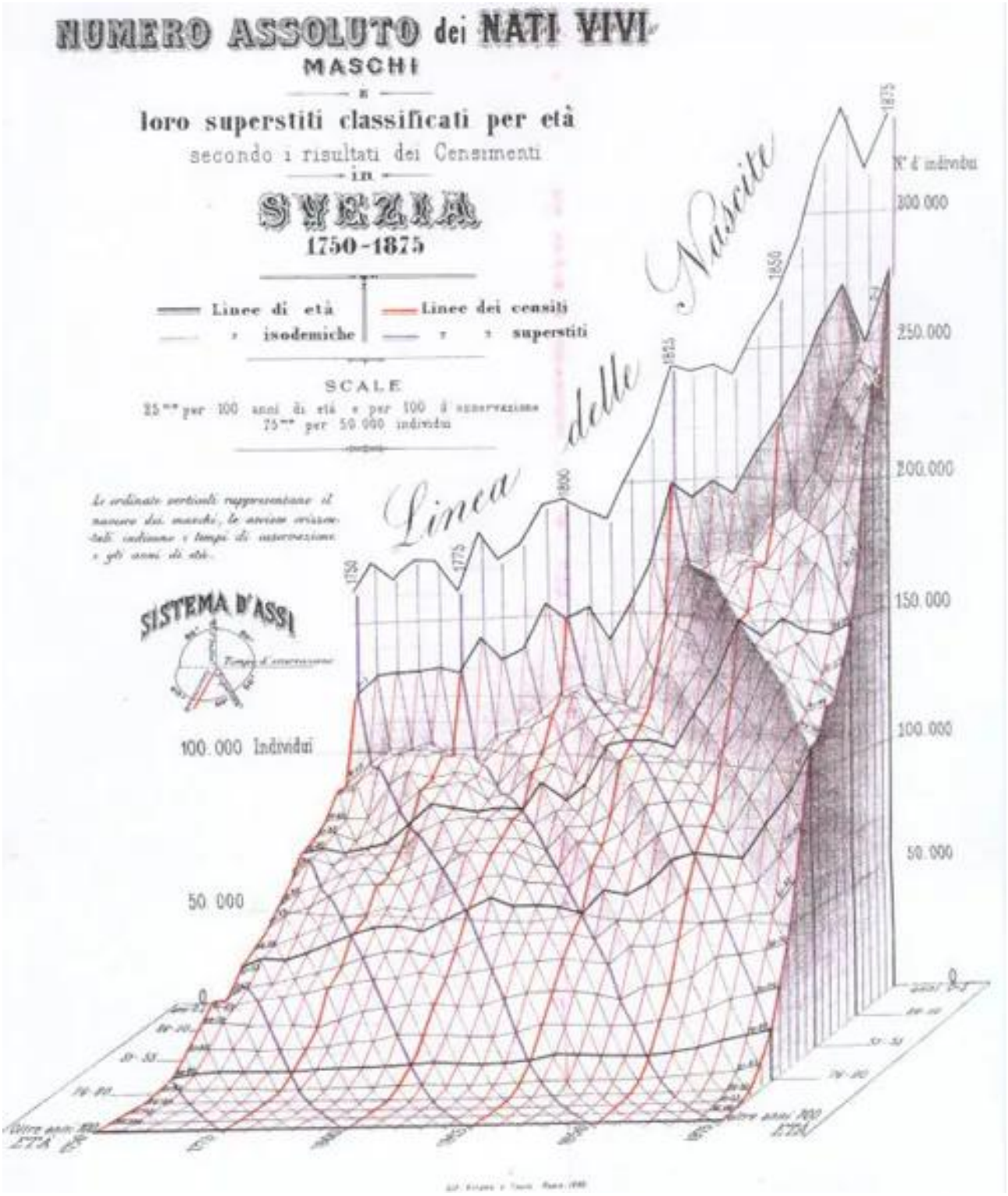
reads time from left to right. The growing trend towards statistical visualizations hit a small roadblock in the early 20th century. Friendly describes this era as the modern dark ages for data visualization. Statisticians were increasingly concerned with exact numbers, and considered images to be overly inaccurate. While innovation in the field may have indeed veered away from data visualizations, this period saw a growth of data visualization in the public consciousness. Charts and graphs of various sorts were rapidly becoming adopted into text books, business applications, science, and government. While statisticians gave data visualization the cold shoulder during this era, the first half of the 20th century also brought psychology forward as a science; in particular, the development of cognitive psychology and the study of human perception has provided a much better understanding as to how the brain interprets information and recognizes patterns. This research has been instrumental in developing and refining the science of data visualization best practices.



Life span of famous persons

Luigi Perozzo, 3D representations

Luigi Perozzo, an Italian statistician, Perozzo produced one of the first 3D representations of data showing the age group of Swedish population between the 18th and 19th centuries. In this diagram years are measured horizontally, numbers of individuals vertically, and age groups (youngest nearest) in depth going into and out of the image. The use of 3D to represent data is now commonplace in contemporary scientific visualization, e.g. medical and engineering sciences.

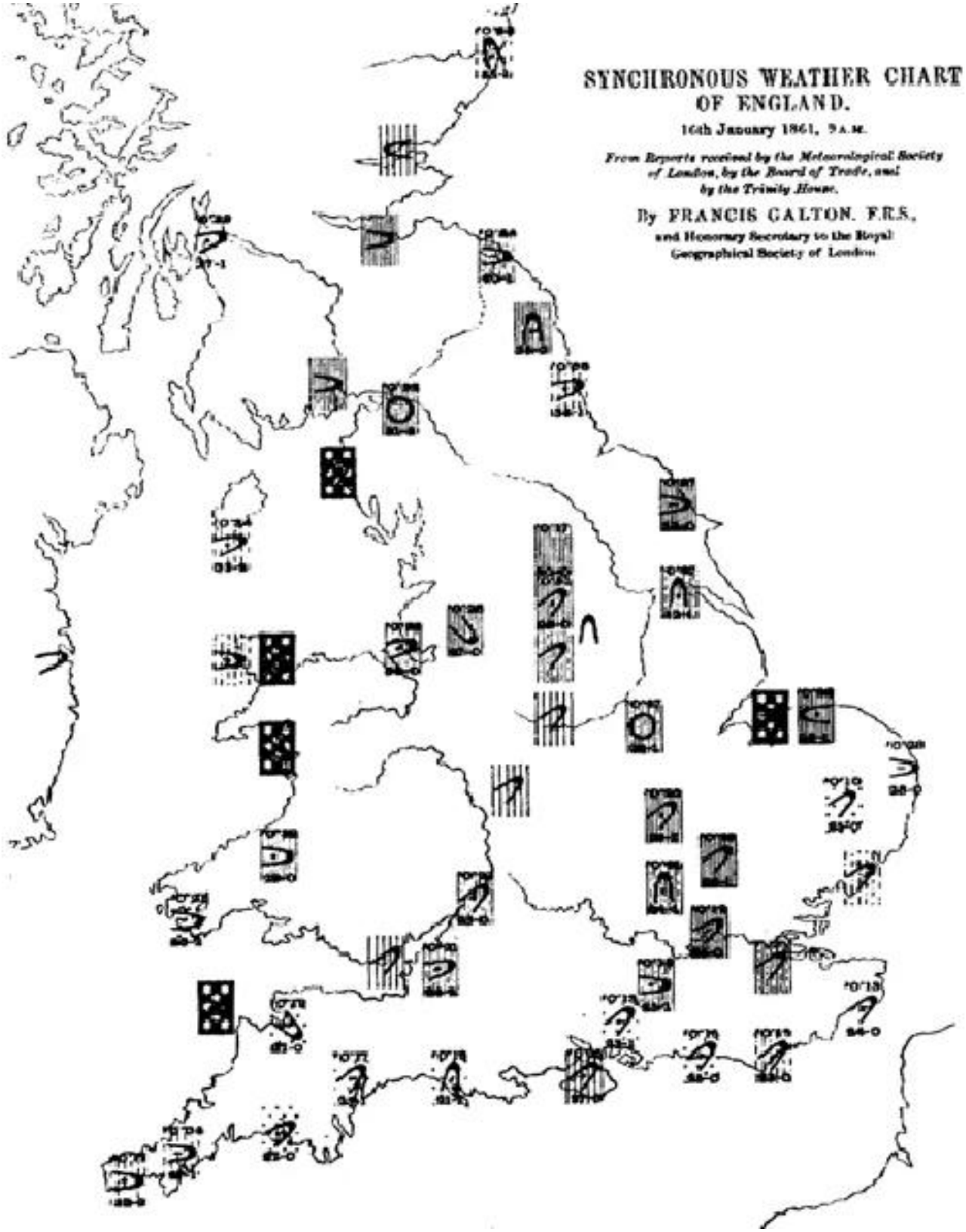


Luigi Perozzo's 3D Representations

Galton, Meteorological Maps

Galton was an English scientist and half cousin of Charles Darwin, his contribution to the history of visualization comes in the form of his meteorological maps and graphical

innovations showing the distribution of air pressure and wind direction. Galton's work paved the way for our modern weather maps.



Galton's Meterological Maps

Charles Booth, Poverty and Crime

Charles Booth's color coded visualizations of London streets, attempted to answer the question of how many Londoners lived in poverty and what caused it. Looking at the map, buildings marked black denote the poorest or in Booth's terms "semi-criminal", dark blue the very poor, purple and pink mixed or people with fairly comfortable incomes, and red and yellow the middle classes and well off. Booth is a pioneer as his work proved the social value of visualizations as a tool for political analysis. His work received much attention from the emerging socially progressive politicians of the day.



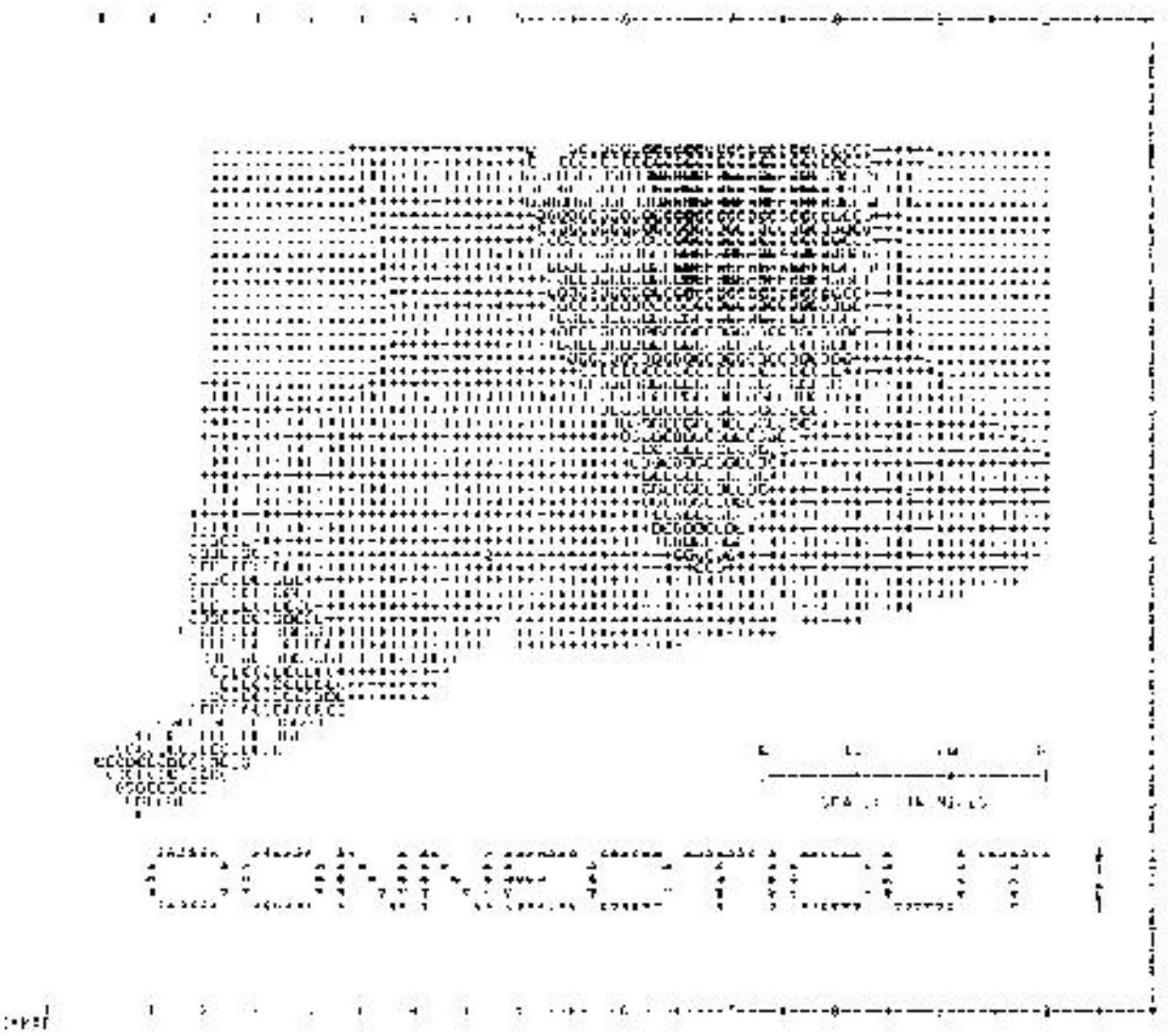
Charles Booth, Poverty and Crime

The latter half of the 20th century is what Friendly calls the rebirth of data visualization, caused by the emergence of computer processing. Computers gave statisticians the ability

to collect and store data in increasingly larger volumes, as well as the ability to visualize the information quickly and easily. By the early to mid part of the 20th century information visualization had entered mainstream use and had become commonplace in magazines, cinemas and newspapers, but it was the emergence of computer technology that was to produce the next revolution in the subject. During the Second World War computers had proved very effective at handling the vast amounts of data needed to analyze military intelligence, providing a post-war technological platform for the development of new ways of plotting statistics graphically. By the 1970s and 80s the first full color computer visualizations were being developed and interactive graphics were providing new ways of revealing the stories hidden in data by enabling navigation of it in three-dimensions. The 1960s and 1970s saw the emergence of researchers like John W. Tukey in the United States and Jacques Bertin in France, who developed the science of information visualization in the areas of statistics and cartography, respectively.

Howard Fisher, Software Breakthroughs

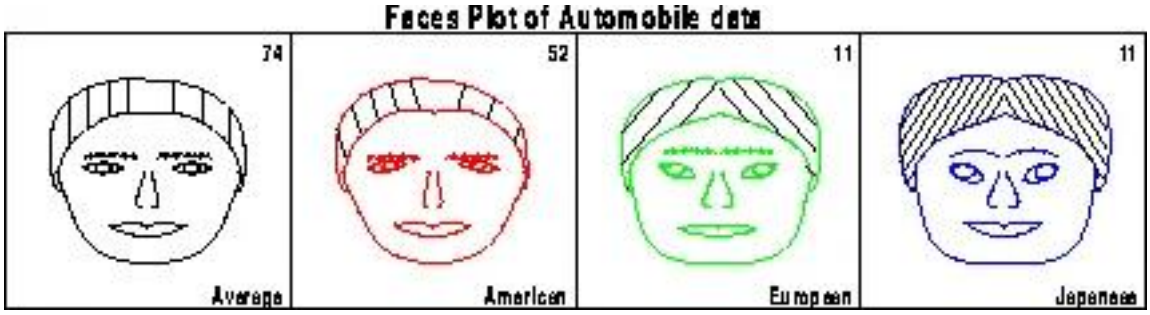
Working at the Harvard Laboratory for Computer Graphics and Spatial Analysis in the 1960s, Fisher developed the first general purpose mapping software.



Howard Fisher, Software Breakthroughs

Herman Chernoff, Chernoff Faces

In 1973 statistician Herman Chernoff used cartoons faces as a way of displaying data. They rely upon our mind's ability to spot changes in human faces, each component of the face representing a different form of data.



Chernoff, Chernoff Faces

Richard A. Becker and William S. Cleveland, Interactive Visualisations

Interactive graphic system from 1987 allowing direct manipulation of data through interaction. The early 80s saw the emergence of Edward Tufte, whose seminal work, The Visual Display of Quantitative Information 6 is still used today in university courses for data visualization and statistical analysis. Tufte also introduced us to the sparkline, which gives the general shape of a trend in a small amount of space. The last two decades have seen the field of data visualization explode into dozens, and even hundreds of focus areas.

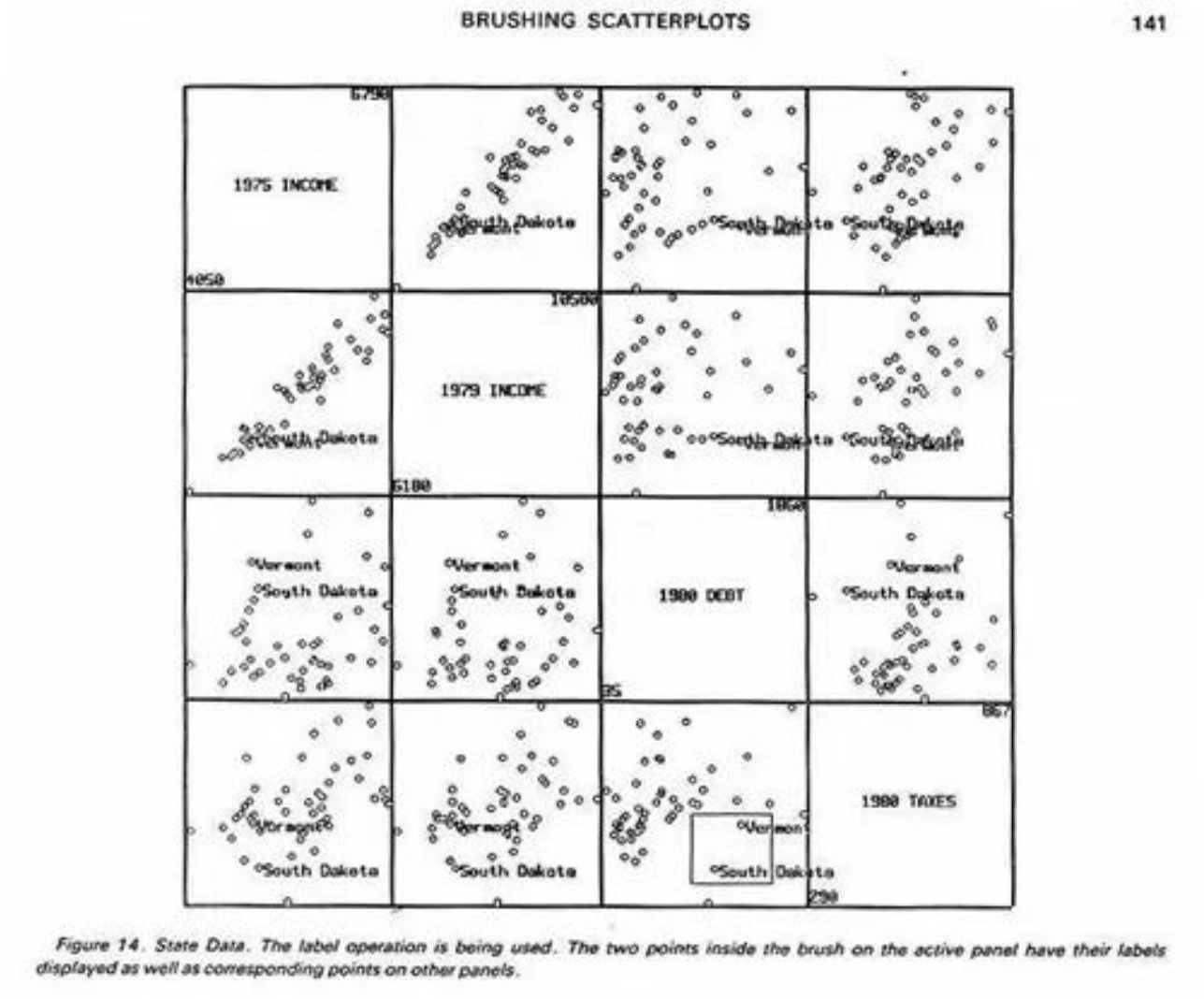
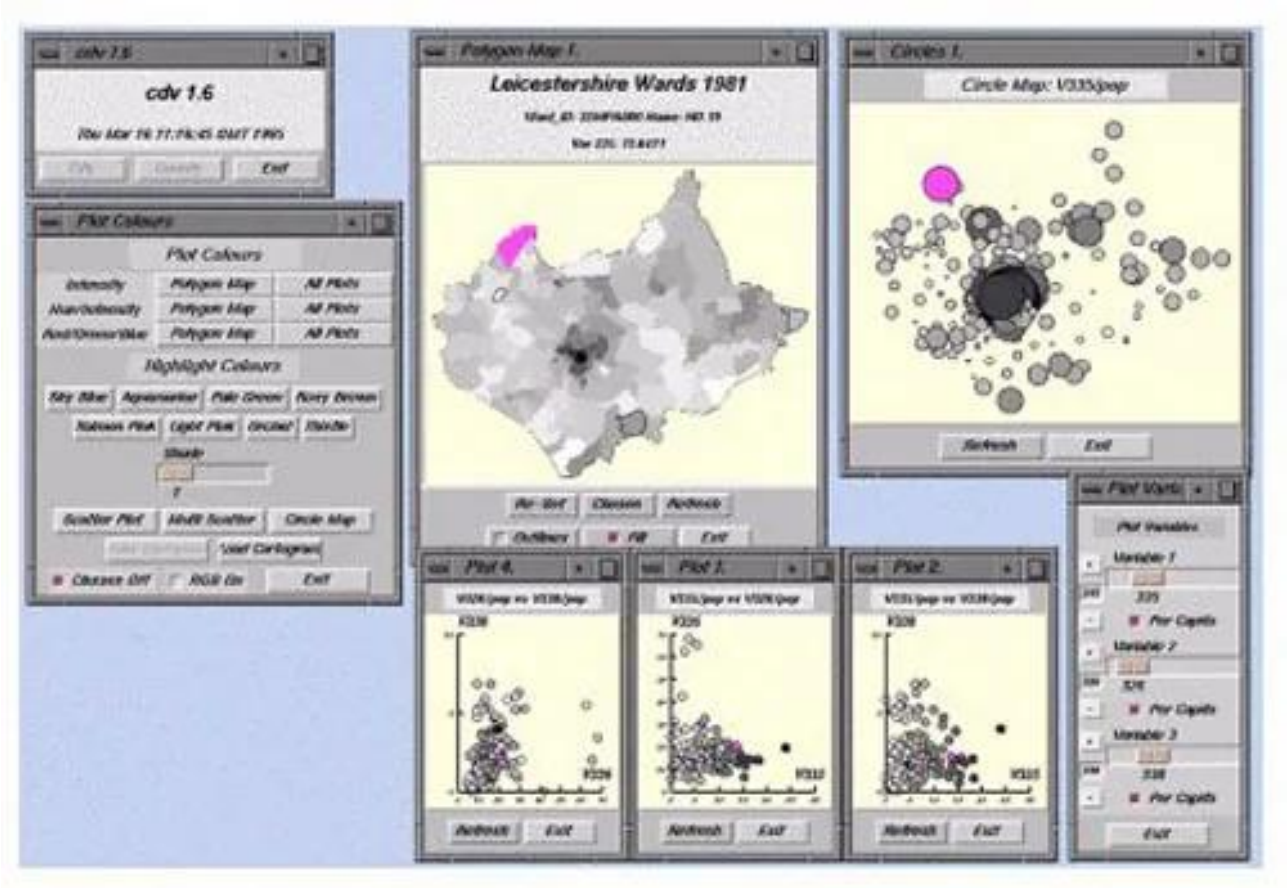


Figure 14. State Data. The label operation is being used. The two points inside the brush on the active panel have their labels displayed as well as corresponding points on other panels.

Cleveland - Interactive Visualisations

Jason Dykes, Cartographic Data Visualiser

Jason Dykes, Cartographic Data Visualiser, developed a map visualisation toolkit in 1996. The toolkit included interactive options for exploring combinations of geography and information. Dashboards and data discovery tools, scorecard applications, analytics suites, and an assortment of other software tools enable businesses, researchers, and individuals to explore their data in new and increasingly imaginative ways. Notable writers and educators in the modern era, including Alberto Cairo, Stephen Few and Colin Ware, continue to refine the science and art of data visualization and bring it to new heights.



Cartographic Visualiser

Summary

The emergence of the Internet in the latter part of the 20th century, the availability of new software tools such as Flash, Google Earth and Processing, and the increase in publicly available data, has seen a huge increase in types of data visualization. Whereas in the past, various graphical aids for interpreting data have mainly been produced by specialist statisticians and scientists, in this new era we see an increasing appetite amongst members of the general public to produce their own. We live in an exciting yet challenging time for data visualization. Even as we discover new ways to collect, aggregate, analyze and visualize data, we are also discovering some new and important societal challenges regarding the violation of privacy and the potential misuse of data – both intentional and unintentional. As we enter the information age, it's both exciting and terrifying to imagine what the future holds in store for us, both as individuals and as a society. Randy Bachman said it best:

We ain't seen nothin' yet!