On History of Information Visualization

Mária Kmeťová
Department of Mathematics, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, Nitra, Slovakia
mkmetova@ukf.sk

Keywords: abstract data, data visualization, distant learning, graphics

Abstract: Information visualization is a new discipline using computer graphics technologies and based on piece of knowledge in statistics, informatics, geometry and psychology. Visualization offers technique for creating images, charts and animations for communicate some information. This is the merit of information visualization doing it extremely important for distant learning. Good and powerful visual representation of abstract data may help in communicating, analyzing data and confirming hypotheses and can substitute many pages of explanatory text. Examples from history give inspirations for building excellent visual representations, but wrong examples should be avoided in next creation of visualization.

Introduction

The history of visualization is that of the search for new artefacts to amplify the ability to know; it is the history of writing and of maps, the history of knowledge (Dürsteler, 2002). Since humans perceive visual attributes very well, we can represent a great deal of different data visually. The graphic representation of abstract data has deep roots reaching into the history of the earliest cartography. The first known city plan was created in the late 7^{th} millennium BCE in Çatal Höyük in Anatolia (modern Turkey) and was painted onto a wall.

Visual Representation of Knowledge

We can find many examples of the developments and innovations related to the visual representation of knowledge. An exemplary use of map to chart patterns of disease was the famous dot map of Dr. John Snow, who plotted the location of deaths from cholera in central London for September 1854 (Tufte, 2001). Examining the map on the Figure 1, Snow observed that cholera had spread among those who lived near and drank from the Broad Street water pump. This is an example, where graphical analysis of data is more efficient as a math calculation.

Competently remake cartography maps can have narrative graphic function too. An example of excellent space-time story graphic is from French engineer, Charles Joseph Minard (1781-1870) who visualized Napoleon´s Russian campaign of 1812 (Figure 2) with combination of data map and time-series.

Figure 1: Snow´s map of location of deaths from cholera in central London in 1854.

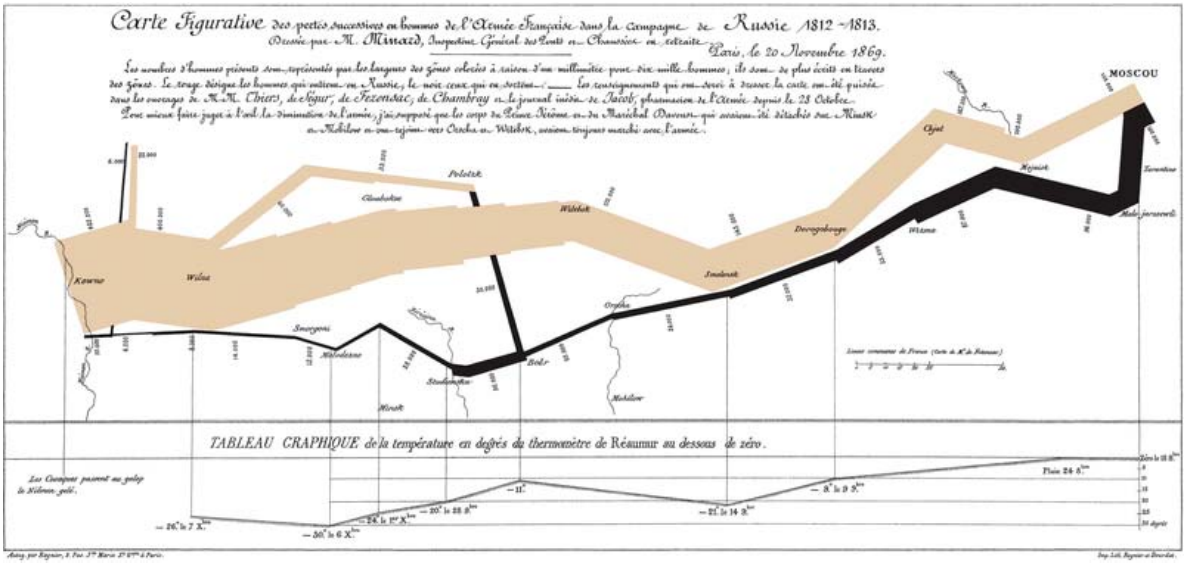

Figure 2: Minard´s visualization of Napoleon´s Russian campaign of 1812, drawn in 1861.

Six variables are plotted: the size of the army, its geographic location (2 variables), direction of the army´s movement, and the temperature on various dates during the retreat from Moscow. This excellent space-time story graphic illustrates, "how multivariate complexity can be subtly integrated into graphical architecture, integrated so gently and unobtrusively that viewers are hardly aware that they are looking into a world of four or five dimensions" (Tufte, 2001).

Knowledge visualization is an important part of information visualization. Visual representations are used very successfully for transfer the knowledge between teacher and students. Both, computer and non-computer based visualization methods can be used in learning process and can substitute many pages of explanatory text. A simple and well-known example of visualization of mathematical proof (cited from (Šedivý et al., 2000)) is shown on the Figure 3. Tufte emphasizes the role of colours in this type of visual representation (Tufte, 1990, Chapter 5: Color and Information).

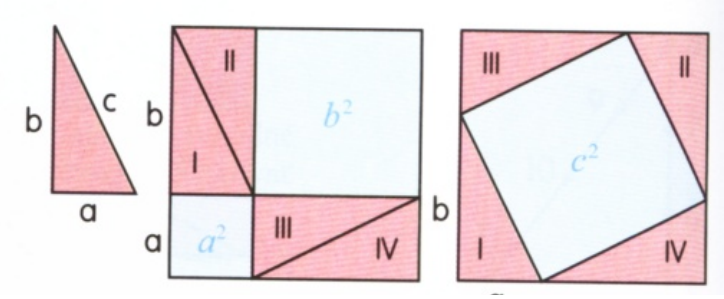

Figure 3: The visual proof of Pythagorean Theorem.

Data Visualization

The first author, who dedicated his book (The Visual Display of Quantitative Information, 1983) to information visualization, was E. R. Tufte, political economist and statistician from Yale and Princeton University and the member of American Statistical Association. He is considered as a founder of this new discipline. The recent emphasis on visualization started in 1987 with the special issue of Computer Graphics on Visualization in Scientific Computing. Since then there have been several conferences and workshops, co-sponsored by the IEEE Computer Society and ACM SIGGRAPH. They have been devoted to the general topics of data visualization, information visualization and scientific visualization, and more specific areas such as volume visualization.

Quantitative Data Visualization

The simplest way to depict and organize quantitative data is a table. However, if the data set has several dimensions, the table became confused. Since the 16th century, techniques for precise observation were well developed and then we can see the effort to show mathematical variables graphically. In the 18th century, new forms of visualization of economics, demography and health data appear. Technological innovations (colour, press) open new possibilities of data representation in printed media. The first half of the 19th century was responsible for an explosion in the growth of statistical graphics. All forms of statistic charts known today were developed at this time (Friendly, 2005). One of the famous economic data visualization is Playfair´s chart (Figure 4), where three parallel time-series are plotted: prices, wages, and the reigns of British kings and queens for 250 years, from 1565 to 1821.

Second half of 19th century is known as the golden age of data graphics (Friendly, 2005). Minard´s flow maps, pie chart, cartogram, polar coordinates chart, three-dimensional stereogram, and modern weather maps arise all from 19th century.

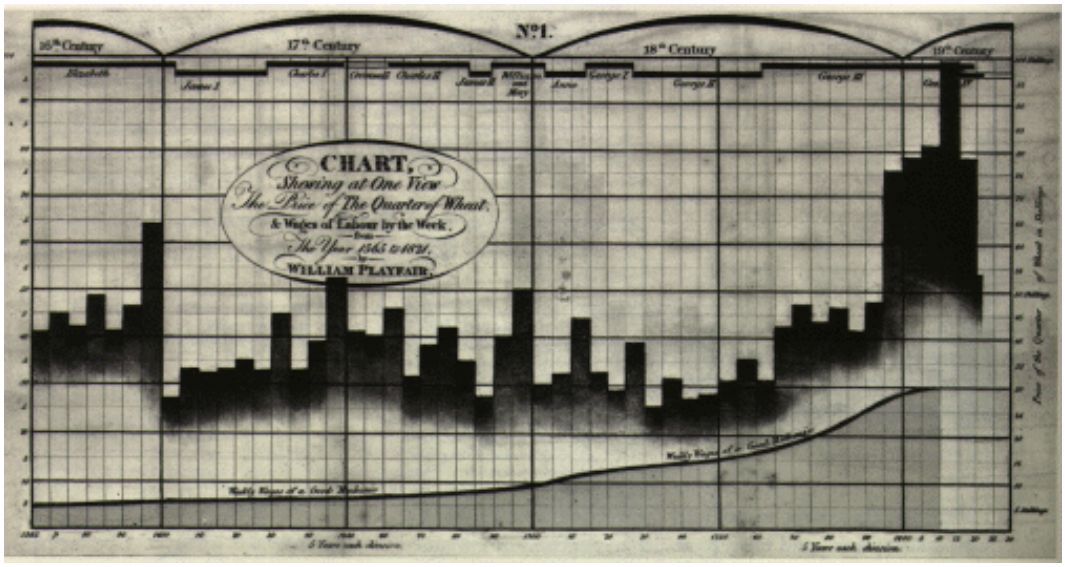

Figure 4: Chart with economic data. William Playfair (1759-1823), England.

The development of software and computer systems brought possibility of visualization of very large multidimensional data sets and manipulation with them. Geometer Alfred Inselberg, researcher at IBM, had the idea of defining a geometric space for representation of multivariate data as an *n*-space through an arbitrary number of axes, arranging them parallel. This was the origin of parallel coordinates, the one of today´s most common technique of visual representation of multivariate data (Mazza, 2008). Parallel coordinates display high-dimensional data points. A traditional, Cartesian scatter-plot of high-dimensional points only shows two or three dimensions at a time. A parallel coordinate view represents each data point as a line that traverses parallel axes mentioned above. Each line crosses each axis at the location determined by the point's value in that dimension. With this view all dimensions are represented at once and they show the correlation between adjacent dimensions.

Ordinal and Categorical Data Visualization

Ordinal and categorical data are not characterized numerically, but they are given in order (not necessary linear, can be organized in a network too) or as a member of category (belongs to). Visualization of this type of data is not so frequent as the visualization of quantitative data, but we know very different visual representations of such data, for instance Venn-diagrams, flow-charts, trees, and so on. An example of a brilliant representation of network data is the Map of the London subway from year 1933, which became a world-recognized symbol for its simplicity and clarity (see Figure 5 cited from http://www.probertencyclopaedia.com).

A new method for representation of mainly categorical data is the so-called parallel sets, a technique developed by a group of researchers from the VRVis (Virtual Reality and Visualization) Institute of Vienna in 2006.

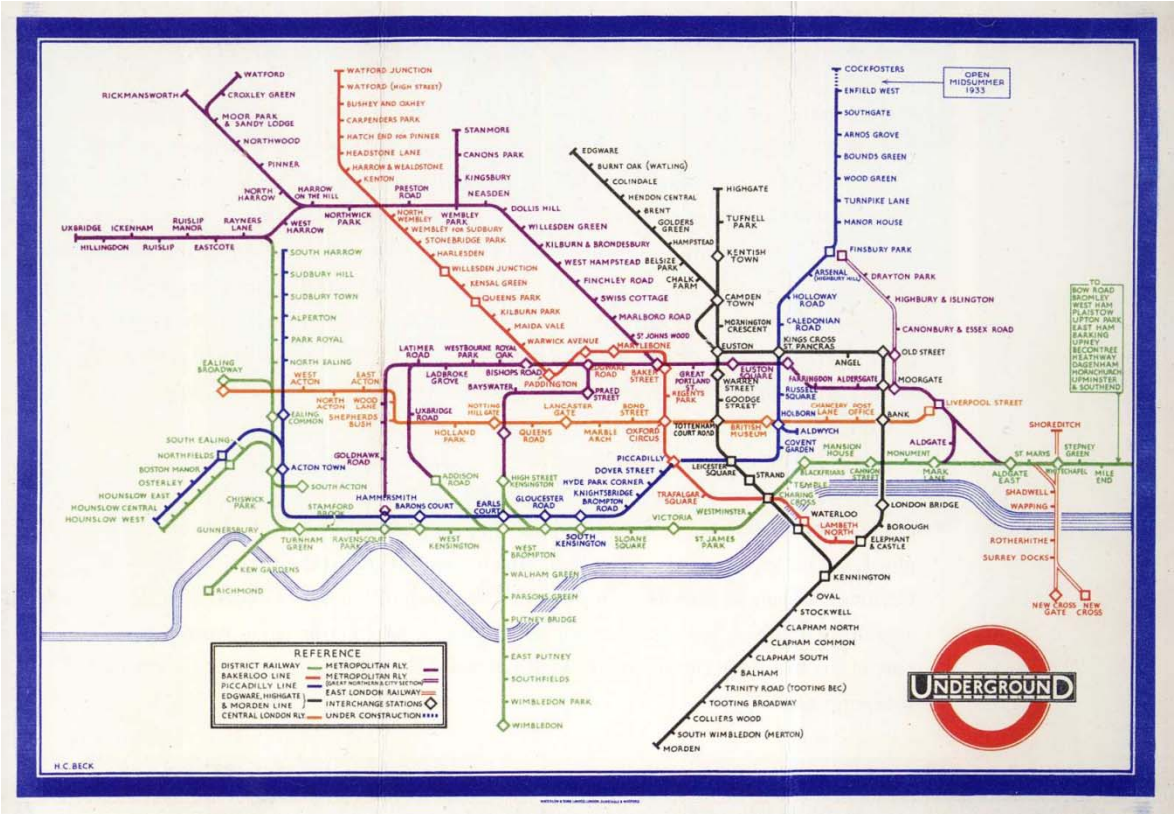

Figure 5: Map of London Underground created by Harry Beck, 1933.

Parallel sets take inspiration from the parallel coordinates but, in contrast, the frequency of the values in the dataset substitutes for the representation of each single instance. This type of representation, unlike the parallel coordinates, turns out to be more appropriate in the case of

categorical data. It can manage very large dataset without the problem of space (Mazza, 2008). Figure 6 shows a representation in parallel sets of a dataset derived from the victims of the Titanic disaster. The attributes arranged next to each other are linked by connections such that they represent the values of frequencies in which the conditions are verified. There are shown 3 categorical attributes of the dataset: the class in which they travelled, the sex, and whether they survived. The figure very clearly shows that most of the first-class passengers survived, while the majority of the third-class passengers and the crew died in the disaster (Mazza, 2008).

Conclusion

The graphic representation of abstract data has deep roots reaching into the history of the earliest cartography, later into economics and statistical graphics, medicine and other fields. Developments in printing, reproduction and practical observation enabled the wider use of graphics along the way. From earliest history there are known town maps and navigation maps mainly. In the 16th century, techniques for precise observation were well developed and then we can see the effort to show ideas and mathematical variables graphically. In the 18th century, new forms of visualization of economics, demography and health data appear.

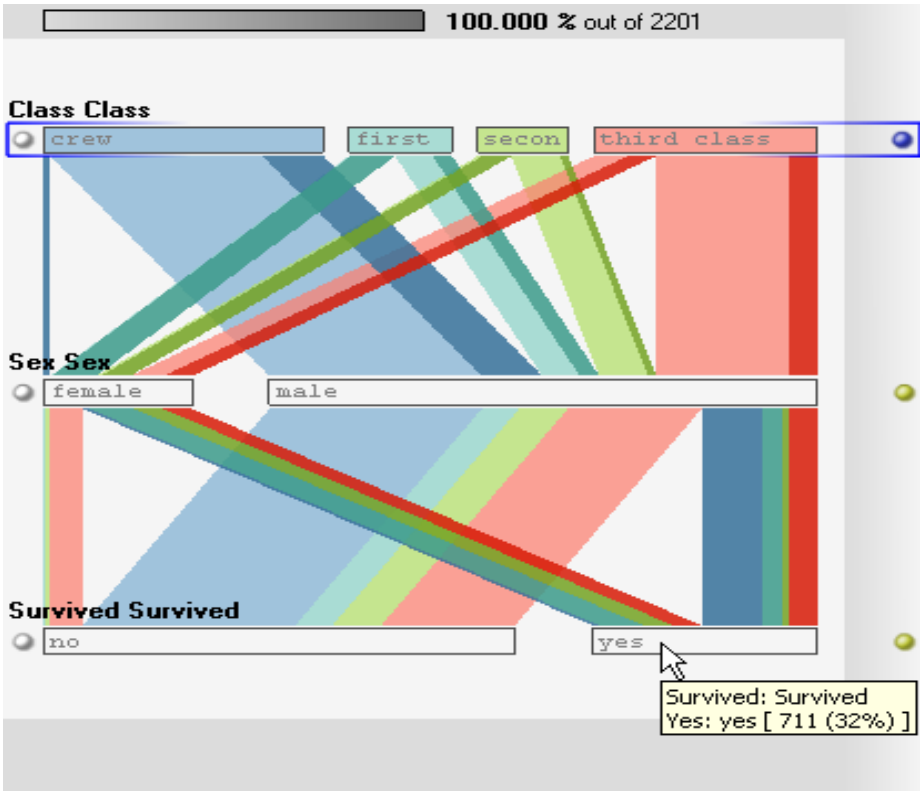

Figure 6: The Titanic disaster visualization by parallel sets (http://srvac.uncc.edu/research/parallelSets.html).

Technological innovations (colour, press) open new possibilities. Explosion in quantum of data implies the beginning of modern infographics in the first half of 19th century. All forms of statistic charts known today were developed at this time. Second half of 19th century is known as the golden age of data graphics. Minardi's illustration of Napoleon's campaign against Russia arises from this time. This is "the best graphic ever made" according to evaluation of E.R. Tufte.

Nowadays, the development of software and computer systems influences the invention of graphic techniques and methods of multidimensional visualization.

References

Friendly, M., 2005. Milestones in the history of data visualization: A case study in statistical historiography. In C. Weihs and W. Gaul, eds., Classification: The Ubiquitous Challenge, (pp. 34–52). New York: Springer.

Mazza, R., 2009. Introduction to Information Visualization, Springer-Verlag London.

Šedivý, O., Čeretková, S., Malperová, M., Bálint, Ľ., 2000. Matematika pre 8. ročník základných škôl, 1. časť, Slovenské pedagogické nakladateľstvo, Bratislava

Tufte, E. R., 2001. The Visual Display of Quantitative Information, 2nd edition, Graphic Press, Cheshire, Connecticut.

Tufte, E. R., 1990. Envisioning Information, Graphic Press, Cheshire, Connecticut.

Tufte, E. R., 1997. Visual Explanations, Images and Quantities, Evidence and Narrative Information, Graphic Press, Cheshire, Connecticut.

Tufte, E.R.,2006. Beautiful Evidence, Graphics Press LLC, Cheshire, Connecticut.

Ware, C., 2004. Information Visualization, Perception for Design, Second Edition, Morgan Kaufmann, Elsevier.

Dürsteler, J. C., 2002. The History of Visualization, Retrieved from http://www.infovis.net

http://www.probertencyclopaedia.com/photolib/maps/Map%20of%20London%20Underground%201933.jpg