

MICHAEL FRIENDLY

DANIEL DENIS

**Discussion and comments. Approche graphique
en analyse des données. The roots and branches
of modern statistical graphics**

Journal de la société française de statistique, tome 141, n° 4 (2000),
p. 51-60

http://www.numdam.org/item?id=JSFS_2000__141_4_51_0

© Société française de statistique, 2000, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

DISCUSSION AND COMMENTS

Approche graphique en analyse des données

The roots and branches of modern statistical graphics

Michael FRIENDLY and Daniel DENIS¹

RÉSUMÉ

Jean-Paul Valois examine le rôle de l'approche graphique dans le développement de diverses méthodes statistiques et propose un système de classification des graphiques. Notre contribution développe ces deux idées en les situant dans un contexte élargi. Nous pensons que, pour appréhender la statistique graphique actuelle, une bonne compréhension de ses racines historiques et de ses branches modernes est essentielle. Nous décrivons ses racines historiques comme un contraste entre résumé numérique et précision d'un côté, exposition visuelle et perception de l'autre. Nous argumentons que ni l'un ni l'autre versant n'est seul suffisant, mais que les deux sont nécessaires pour comprendre la signification des données. Parallèlement, en classifiant les graphiques modernes, il est utile de les envisager selon une typologie de « bas en haut » ou un schéma de « haut en bas ». On conclut qu'il est besoin, à la fois, de la connaissance des fonctions des symboles graphiques et leur compréhension perceptive, comme de la compréhension du but communicatif, pour déterminer si un graphique transmet le message espéré.

ABSTRACT

Jean-Paul Valois examines the role of graphical approaches in the development of a variety of statistical methods and proposes a classification system for the representation of graphics. Our review expands on these two themes, by placing them in a wider context. We argue that in order to understand modern statistical graphics, an understanding of its historical roots and modern branches is essential. We describe its historical roots as a contrast between numerical summarization and precision on the one hand, and visual exposure and perception on the other. We argue that neither one is sufficient on its own, but rather both are necessary to understand the meaning of a data set. Similarly, in classifying modern graphics, it is useful to view them in terms of a bottom-up typology, or a top-down schema. Again, it is concluded that we need both a knowledge of the functions of graphic symbols and their perceptual comprehension, and an understanding of the communication goal to determine whether a graph conveys the intended message.

1. York University, Psychology Department, 4700 Keele Street, Toronto, Ontario, M3J 1P3, friendly@yorku.ca, dand@yorku.ca

DISCUSSION AND COMMENTS

We are grateful to Jean-Paul Valois for his insightful and provocative paper on the role of graphical methods in data analysis, and to the editor for this opportunity to contribute some observations on the issues raised. We believe we can most usefully contribute to this topic by placing these issues in a wider context. So, what follows are a few "supplementary points" to be added to "l'approche graphique en analyse des données".

In our reading, the major themes presented (a) draw links between the development of graphical techniques and cultural history, (b) examine the role of graphical approaches in the development of a variety of statistical methods, and (c) propose a classification system for various graphic representations. Our commentary below relates mainly to (a) the roots, and (c) the branches.

1. ROOTS

With a few scattered exceptions, the earliest seeds of statistical graphics arose first in simple geometric diagrams, then in the making of maps, with relatively precise longitude and latitude, to aid in navigation and exploration. Later seeds were planted in the use of coordinate systems to depict more complex, but still functional relations among variables. Only in the late 17th and 18th centuries, did ideas about errors of measurement, probability theory, and the beginnings of demographic statistics, create a fertile climate for the use of charts and diagrams to portray empirical data. Over the 18th and 19th centuries, numbers pertaining to people and states – social, moral, medical, and economic statistics began to be gathered in large, periodic, and systematic series. At the same time, the usefulness of such data, for business, for government, and as a subject for study in its own right, began to be recognized.

This sprouting of statistical thinking was accompanied by a rise in visual thinking. For example, nomograms were developed to perform calculations geometrically (Hankins, 1999); new graphic forms were invented to make the properties of empirical data more easily accessible to visual inspection; and the close relations between the official numbers of the state (the origin of the word "statistics") and its geography gave rise to the visual representation of such data on maps, now called "thematic cartography" (Palsky, 1996). These early roots of modern statistical graphics are documented in our project on "Milestones in the History of Thematic Cartography, Statistical Graphics, and Data Visualization," (Friendly & Denis, 2001).

1.1 Early sprouts

In the historiography of any field, it is always the abrupt changes – the empty gaps, and the great leaps forward, that attract attention, and require explanation. Valois notes the relation (or contrast) between statistical graphics and the developments in mathematics and applied statistics. Specifically, he points to the historic, non-parallel use of graphics versus pure statistical

DISCUSSION AND COMMENTS

numerical solutions – the conflict between visual and the formal approaches to data.

For instance, he cites Descartes' claim, echoed by Malebranche (1674), that "our eyes generally deceive us" but that pure mathematics might "correct" for this as one example of how mathematical precision was used not in parallel, but rather *in place* of graphical methods. Valois continues that it was not until Locke (1693) and the empiricists that confidence in visual perception was restored. But after Descartes, it took a century and a half until William Playfair (1786) introduced a wide variety of new graphic forms, and claimed to make "the numbers speak to the eyes." It is this relation, between summarization – the numerical precision of statistics, and exposure – the visual, applied use of statistical graphics, that we wish to expand on and explore in this review.

1.2 Modern sprouts

Another gap, followed by a great leap forward occurred in the 20th century. The late 19th century was an "age of enthusiasm" for statistical graphics (Palsky, 1996), which saw the production of the most varied and beautiful statistical charts, diagrams, and maps ever produced. In contrast, the period from the mid-1930s until the mid-1960s might be called the "modern dark ages" of statistical graphics.

By the late 1930s, the enthusiasm for visualization had vanished and would not reappear until the 1960s-1970s. This new enthusiasm was largely a product of the seminal work of Tukey in the U.S. and Bertin in France. Here, we explore the following questions : 1) why was there a lapse of interest in statistical graphics in the middle of the 21st century? 2) why were Tukey and Bertin's graphical approaches welcomed so enthusiastically following this lapse? 3) what does this finding say about the relation between statistical graphics and the employment of mathematical statistics? Surely, these questions are interrelated. For instance, to learn why Tukey's and Bertin's methods were successful, one needs to first know why the lapse occurred. Knowing this might shed light on why there was an "explosion" of graphical techniques following the void.

1.2.1 *Death by precision*

The most plausible explanation for the lapse in interest in statistical graphics is the rise of quantification and statistical models, which began perhaps with Galton, and was developed largely by Pearson. By the early 20th century, statistical methods had escaped the two-dimensional plane, with the initial development of factor analysis by Spearman (1904), principal components analysis by Hotelling (1933), and the related contribution by Eckart & Young (1937) of low-rank approximations of a matrix. These ideas would later provide fertile ground for data visualization. We do not deny that Pearson and others used occasional graphical displays (Stigler, 1986, p. 327). But, at this time, quantification and precision captured the imagination of statisticians and research workers.

DISCUSSION AND COMMENTS

For illustration, we focus here on the impact of Fisher's (1925) landmark publication of *Statistical Methods for Research Workers*. Fisher aimed to give the experimental scientist a precise, mathematical method for quantifying uncertainty – for going from the observed data at hand, to an inference about the populations from which the data were drawn : "We may at once admit that any inference from the particular to the general must be attended with some degree of uncertainty, but this is not the same as to admit that such inference cannot be absolutely rigorous, for the nature and degree of the uncertainty may itself be capable of rigorous expression" (Fisher, 1925, p. 4). As noted by Lovie (1981), the use of the ANOVA model became widespread only 3-4 years after Fisher wrote *Statistical Methods*.

Thus, the precision provided by Fisher's techniques in experimental design (and other developments in statistics) may in part account for why statistical graphics innovations were practically non-existent soon after his contribution. After all, why graph data, when you can calculate F-ratios and p-values that provide firm conclusions, with known probabilities of error? It may have been this emphasis on conclusions and confirmatory analyses, rather than visualization, which led to the lapse of graphical innovations in mid-century.

1.2.2 *Re-birth*

Suppose we provisionally accept the theory that the statistical precision behind the innovations of Fisher and others (Pearson, Gossett) had an influence on the demise of graphical innovation in the middle of the 20th century. Our task then becomes one of explaining why there was a resurrection of graphics following this dormant period.

In the U.S., we attribute to Tukey the successful rebirth of interest in data analysis (as distinct from statistics), graphical methods, and the distinction between exploratory and confirmatory modes of analysis. "Once upon a time, statisticians only explored. Then they learned to confirm exactly... The connection of the most used techniques with past insights was weakened. Anything to which a confirmatory procedure was not explicitly attached was decried as 'mere descriptive statistics', no matter how much we had learned from it." (Tukey, 1977, p. vii). In France, the honors must be shared between Bertin, who led to the revival of interest in graphics with the general theory of graphic signs, their combinations and meaning embodied in the Semiology (Bertin, 1967), and Benzécri (1973), who established l'analyse des données as an important topic, worthy of development, with a strong emphasis on graphical presentation and description.

We will entertain two possibilities for these successes : 1) the advance in computer technology provided a means for easy graphing, and 2) a Kuhnian-type paradigm shift, brought about by the realization that mathematical methods, precise as they may be, are not in themselves adequate for complete and thorough data analysis.

The first possibility, of computer technology, Beniger and Robyn (1977) argue contributed to the rise of statistical graphics. "All of these innovations

DISCUSSION AND COMMENTS

[including Tukey's], exploit computer technology, and most have little practical value except when executed by a computer (p. 7)." In the foreword to Bertin (1983), Wainer claims that the massive amount of available data, combined with the technology of the computer, is responsible for the rise in graphics : "This graphic explosion, though caused by the need to present massive amounts of information compactly, is abetted by the computer (p. vii)". We find this claim plausible, but unconvincing.

Despite the present ease use of graphics software, many of Tukey's and Bertin's innovations were explicitly designed for manual use. As Tukey says in (*EDA Exploratory Data Analysis* (1977), "everything illustrated or set as a problem can be done with pencil and paper... the only tools the illustrator used were a pen and a straightedge (p. x)". Computer technology and sophisticated display devices have surely contributed to the rapid re-growth of graphics since the late 1970s, but they were neither a prerequisite nor the causative factors. The invention of the toaster – a new way to burn bread – would not have been popular unless people wanted to eat toast. Similarly, without an impetus for graphics, the advent of computing might only have led to better ways to calculate p -values. Finally, as Friendly (2000) has noted, many of the most beautiful and sophisticated graphics ever produced were contrived long before the first computer, which suggests that computer technology played only a *part* in the modern rise of graphics.

The second possibility, that of a paradigm shift (Kuhn, 1962) is well worth exploring¹. We suggested earlier that the rise of mathematical statistics (typified by Fisher) had dampened the interest in graphical approaches to analyzing data. Yet, after 30 years of developments and practice, chinks in the armor of what we now call "classical inference" had begun to appear. What happens to our p -values when assumptions – the Holy Trinity of normality, constant variance, and independence – are not met? We speak here to the effect of Tukey's contributions; Palsky and Robic (1997) provide an account of the development of Bertin's ideas and their influence.

Beginning in the mid-1960s, Tukey began to announce a new philosophy of **data analysis**, related to, but distinct from statistics. It is hard to date this philosophy exactly, but "The future of data analysis" (Tukey, 1962) is the earliest clear statement. "For a long time I have thought I was a statistician, interested in inferences from the particular to the general. But as I have watched mathematical statistics evolve, I have had cause to wonder and to doubt. ... All in all I have come to feel that my central interest is *in data analysis...*" Over the next 15 years, until the eventual publication of *EDA*, Tukey published 108 papers, nearly 8 per year. We believe it is the scope and breadth of this philosophy, as much as the simplicity and utility of specific graphical techniques (stem-leaf plots, boxplots, hanging rootograms, etc.) which are responsible for the paradigm shift towards *EDA* and the rebirth of

1. To avoid argument with historians of science, we admit to using the term "paradigm" loosely, but so does Kuhn (Masterman, 1970). We refer here to how the usual practice of statistics and graphics were fundamentally altered by Tukey's philosophy of data analysis and techniques, and Bertin's classification of graphical signs and meanings.

DISCUSSION AND COMMENTS

graphics. Tukey's emphasis on the complementary roles of confirmation and exploration, of numerical summarization and graphical exposure, did much to establish the importance of this philosophy of data analysis : "We can no longer get along without confirmatory data analysis. *But we need not start with it.*" (Tukey, 1973, p. vii)

We end this section with a personal observation. The first author began his career in quantitative psychology, with application to cognition and human memory. At that time (1965-70) it was nearly impossible to publish in experimental psychology journals without an ANOVA table with p -values, decorated by appropriate sets of *s, proclaiming statistical significance. Graphs were sometimes accepted by journal editors, but begrudgingly, and largely as decoration. The real proof was in the p -values. Tukey changed things, not only in statistics, but in fields like experimental psychology as well. He "put an end to the view that graphics were only for decorating a few numbers" (Tufté, 1983, p. 53). A recent testament to the impact of this paradigm-shift on "normal science" is the editorial statement by Loftus (1993) that "a picture is worth a thousand p -values" and in many cases can completely replace numerical analysis of sample statistics. Tukey never went as far; but such is the nature of paradigm shifts.

2. BRANCHES

Schemes of classification for the plantings in any discipline are intended to organize the observed species in a coherent structure, to facilitate theory-based comparisons, and to provide fertile soil for the husbandry of novel ideas and new species. Throughout history, there have been two opposite approaches to these questions – the bottom-up, empiricist version which stems from Aristotle, through Locke, and seeks to establish general laws and regularities through observed similarities, and the top-down, rationalist version, which stems from Plato through Descartes, and seeks to demonstrate general laws through pure reason.

In our metaphor of growth, the difference, roughly stated, is whether you look first at the leaves and flowers, or at the trunk and it's main branches, to account for the diversity and health of the species. A more productive analogy, we will claim, is with linguistic analysis, where bottom-up approaches to phonology, syntax, and semantics, may be contrasted with top-down approaches which place greatest emphasis on the pragmatics and meaning of communicative acts.

2.1 Bottom-up

In statistical graphics, most typologies and classifications follow the bottom-up, empiricist tradition, and regard a graphic display as a "stimulus", a packet of information to be conveyed to an ideal observer. Bertin's (1967) *Sémiologie graphique* still stands as the landmark analysis of graphic signs and symbols.

DISCUSSION AND COMMENTS

Bertin's analysis is exquisitely detailed in terms of "implantation" (point, line, area), properties and uses of "retinal variables" (size, value, texture, color, orientation, shape), combined with an analysis of the variables represented as quantitative (Q), ordered classes (O), categorical, non-equivalence classes (\neq) and so forth.

Valois expands admirably on this bottom-up scheme, incorporating dimensions of the number of variables, type of question asked, and coordinate system (Cartesian, radial, parallel) used in the graphic representation. In a similar way, Cleveland (1993) proposed a model to account for the accuracy of visual decoding of graphics in terms of classifications of the information presented (Q, \neq), visual processing required (pattern perception vs. lookup), and the visual operations required for these tasks. Yet, Cleveland's ranking of graphical elements is based largely on one type of task : magnitude estimation. When the viewer's task is different (detecting clusters, comparing part to whole, etc.) the ranking of visual attributes differs. Hence, we are pleased to see that Valois includes the type of question asked as a principal organizing dimension.

Perhaps the most comprehensive bottom-up model is Wilkinson's (1999) *Grammar of Graphics*. He proposes a comprehensive, formal analysis of **data** (data sets, variables, attributes, transformations, etc.) **graphs**, organized by both data and geometry (relations, summaries, partitions, networks) into components (point, line, area, contour, path, etc.), and **visual attributes** (form : position, size, shape, etc. ; surface : color, texture, etc. ; motion ; text and so forth). This formal analysis is enriched by an algebra of operators (cross, blend, nest) on data, and a specification language for composing graphs from these elements and operators.

We embrace these bottom-up approaches, but we need something more.

2.2 Top-down

So, what has been forgotten (or ignored) in these bottom-up approaches? If we shift our perspective from the object – the graph, to the subject – the graph-maker, and the graph-viewer, then different criteria for classification become immediately apparent.

Suppose, for example, that we regard a graph as an act of visual communication, analogous to verbal communication (Friendly, 1999). This perspective places the greatest emphasis on the desired communication goal (what do you want to say with a graph?) and the intended audience (yourself, for understanding your data; colleagues, for explaining your findings; the general public, for illustrating their importance). It judges the effectiveness of a graphic by how well that goal is achieved, rather than by accuracy of decoding of elements as in the bottom-up approach. Kosslyn (1985, 1989) has articulated this perspective from a cognitive perspective.

In this view, an effective graphical display, like good writing, communicates ideas with clarity, precision, and efficiency. It is most successful when the message and the graphic design are tailored to the intended audience. Figure 1 shows one organization of visualization methods in terms of primary use, presentation goal, and suggested corresponding design principles. At the

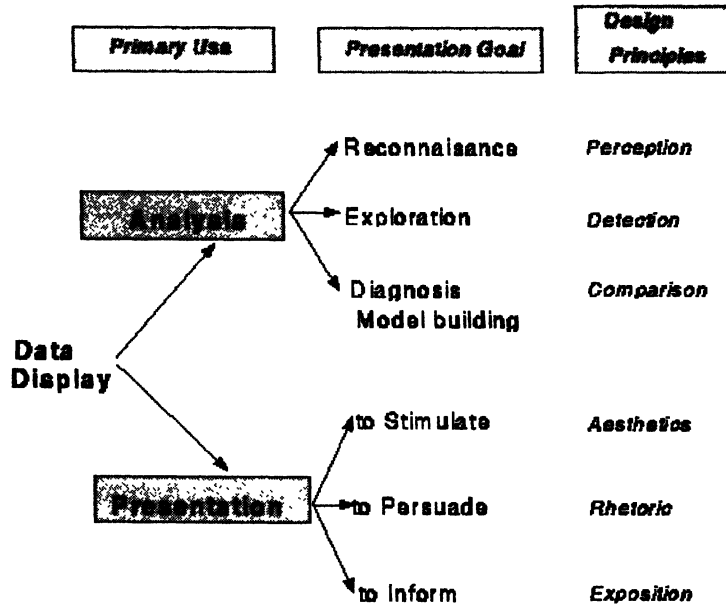
Basic Functions of Data Display

FIG 1. — The basic functions of data display from a top-down perspective.

top level, we distinguish between Analysis and Presentation as the primary communication goal.

Among graphs intended for data analysis, we may distinguish among those designed for (a) **reconnaissance** : a preliminary examination or overview of data (e.g., a scatterplot matrix); (b) **exploration** : graphs designed to help detect patterns or unusual circumstances, or to suggest hypotheses, analyses or models (e.g., boxplots); and (c) **diagnosis** : graphs designed to summarize or critique a statistical summary or model (e.g., influence plots).

Similarly, presentation graphics may be designed to (a) stimulate interest in a topic or phenomenon; (b) **persuade** the reader on some point of argument; or (c) **inform** the viewer about occurrences or relations. In each of these cases, different design principles are most applicable.

As Tukey said, "A picture is not merely worth a thousand words, it is much more likely to be scrutinized, than words are to be read." (Tukey and Wilk, 1965). But, it is equally vital to consider which set of 1000 words a picture might mean.

3. SUMMARY

To understand modern statistical graphics, we need to understand its historical roots and its modern branches.

The historical roots can be understood, perhaps somewhat simplistically, as a contrast between numerical summarization and precision on the one hand, and visual exposure and perception on the other. Neither one is complete. We need *both* visual, exploratory *and* numerical, confirmatory analyses to understand the message contained in any set of data, and to tell a convincing story about its meaning and importance.

Similarly, attempts to classify the modern forms and uses of statistical graphics may be viewed in terms of a bottom-up typology of graphic elements and their combinations, or in terms of a top-down schema focussing on the communication goal of a given display. Again, neither one is complete. We need both knowledge of the functions of graphic symbols and their perceptual comprehension, and an understanding of the communication goal or the viewer's task, to determine if a table or graph conveys the intended message.

REFERENCES

- BENIGER J. R., & ROBYN D. L. (1977). Quantitative graphics in statistics : A brief history. *The American Statistician*, **32**, 1-11.
- BENZÉCRI J. P. (1973). *L'analyse des données*. Dunod : Paris.
- BERTIN J. (1983). *Semiology of graphics*. [Translation by Berg. W.]. University of Wisconsin Press : Madison, WI.
- CLEVELAND W. S. (1993). A model for studying display methods of statistical graphics. *Journal of Computational and Graphical Statistics*, **2**, 323-343.
- ECKART C. & YOUNG G. (1936). Approximation of one matrix by another of lower rank. *Psychometrika*, **1**, 211-218.
- FISHER R. A. (1935). *The design of experiments*. Hafner : NY.
- FRIENDLY M. (1999). Visualizing categorical data. In M. Sirkin, D. Herrmann., S. Schecter, N. Schwarz, J. Tanur, and R. Tourangeau (Eds.), *Cognition and Survey Research*, Wiley : NY., Chapter 20.
- FRIENDLY M. (2000). Re-Visions of Minard. *Statistical Graphics and Computing Newsletter*, **11** (1), 1, 13-19.
- FRIENDLY M. & DENIS D. (2001). Milestones in the history of thematic cartography, statistical graphics, and data visualization.
Web document : [http ://www.math.yorku.ca/SCS/Gallery/milestone/](http://www.math.yorku.ca/SCS/Gallery/milestone/).
- HANKINS T. L. (1999). Blood, dirt, and nomograms : a particular history of graphs. *Isis*, **90**, 50-80.
- HOTELLING H. (1933) Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, **24** : 417-441, 498-520.
- KOSSLYN S. M. (1985). Graphics and human information processing : a review of five books. *Journal of the American Statistical Association*, **80** : 499-512.
- KOSSLYN S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, **3** : 185-225.

DISCUSSION AND COMMENTS

- KUHN T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press : Chicago, IL.
- HOTELLING H. (1933) Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24 : 417-441, 498-520.
- LOCKE J. (1693), *Essai philosophique...* See : complete citation by Valois.
- LOFTUS G. R. (1993). A picture is worth a thousand p values. On the irrelevance of hypothesis testing in the microcomputer age. *Behavioral Research Methods, Instruments, and Computers*, 25, 250-256.
- LOVIE A. D. (1981). On the early history of ANOVA in the analysis of repeated measure designs in psychology. *British Journal of Mathematical and Statistical Psychology*, 34, 1-15.
- MALEBRANCHE N. (1674). *De la recherche de la vérité*. See : complete citation by Valois.
- MASTERMAN M. (1970). The nature of a paradigm. In Lakatos, I. & Musgrave, A., *Criticism and the growth of knowledge*. University Press : Cambridge.
- PALSKY G. (1996). *Des chiffres et des cartes : Naissance et développement de la cartographie quantitative française au XIX^e siècle*. CTHS : Paris.
- PALSKY G. & ROBIC M.-C. (1997). *Aux sources de la semiologie graphique*. In "30 ans de semiologie graphique".
Web document : <http://www.cybergeopresse.fr/semiogr/robig/robigpal.htm>.
- PLAYFAIR W. (1786). *The commercial and political atlas*. Carry : London.
- SPEARMAN C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- STIGLER S. M. (1986). *The history of statistics : The measurement of uncertainty before 1900*. Belknap Press of Harvard University : Cambridge.
- TUFTE E. R. (1983). *The visual display of quantitative information*. Graphics Press : Cheshire : CT.
- TUKEY J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics*, 33, 1-67 and 812.
- TUKEY J. W. (1977). *Exploratory data analysis*. Addison-Wesley : NY.
- TUKEY J. W. & WILK M. B. (1965). Data analysis and statistics : Techniques and approaches. *Proceedings of the Symposium on Information Processing in Sight and Sensory Systems*. (P. W. Nye, ed.), 7-27. California Institute of Technology : Pasadena, CA.
- WILKINSON L. (1999). *The grammar of graphics*. Springer-Verlag : NY.