Journal of Statistical Software

January 2007, Volume 17, Book Review 3. http://www.jstatsoft.org/

Reviewer: Nicholas J. Cox
Durham University

The Grammar of Graphics

Leland Wilkinson
Springer-Verlag, New York, 2005.
ISBN 0-387-24544-8. xviii + 690 pp. USD 79.95.
http://www.spss.com/research/wilkinson/TheGrammarOfGraphics/GOG.html

---

Statistical graphics is, or should be, an transdisciplinary field informed by scientific, statistical, computing, aesthetic, psychological and sociological considerations. The sciences (natural, applied or social and indeed also various literary, linguistic and historical disciplines) generate data relevant to various problems that are amenable to graphical display. Statistics advises or instructs on what should be plotted. Software and hardware are needed to produce displays in practice: very few would now reach for their pens and graph paper as was still common even some twenty or thirty years ago. We clearly should ponder how far graphics are attractive or not, and especially how far graphics work or not, in terms of being interpreted appropriately by users.

But sociological considerations also? I suggest that a full appreciation of the situation of graphics within statistical science depends on recognising and trying to explain a variety of attitudes and behaviours. 'Statistical science' is understood here to range from the science of statistics to any part of science done statistically. Graphics is prominent in almost every introductory course or text and central to a large part of statistical practice. Indeed, I regularly attend scientific presentations largely innocent of p values, t statistics, confidence intervals or even figures of merit such as R^2 or AIC, but which include curve or model fits kept honest, mostly, by intense graphical analysis, large datasets and attention to the underlying science.

On the other hand, in what sense is graphics central to modern statistical research? Or what are we to make of those statistically-minded disciplines currently obsessed by the search for ever more complicated models, to be summarized in terms of endless tables of coefficients and other statistical paraphernalia, and often with scarcely a graph in sight, whether of original data, model results or model diagnostics? To name names would be indelicate, but if the cap fits. . ..

Enthusiasts for graphics can often detect lingering prejudices against their use, even among statisticians who may profess the opposite in public. The old sneer that graphics is for showing the obvious to the ignorant has echoes in some remaining attitudes. The notion that graphics

is limited by subjectivity of interpretation is matched by the still persistent view that no result is worth taking seriously unless it comes with low p value attached or, worse, at least one star against it showing significance at 5%. This notion remains the more firmly entrenched the less practitioners realise how unreliable those indicators really are. And, of course, many statisticians follow mathematical styles and values in preferring the pursuit of new theoretical results to the production of new techniques useful in data analysis.

Here is yet another large question: Why are there so few really good books on statistical graphics? All the best books fit within one foot of shelf space, something that could hardly be said for any other large area of statistics. One likely reason is that only a polymath could do justice to the multi-faceted field that is statistical graphics.

In my view, the Golden Age of statistical graphics (nostalgia's not what it used to be) was a decade around 1980, in which a small cluster of books by some highly talented individuals re-shaped the field: *Exploratory Data Analysis* (Tukey 1977), *Graphical Methods for Data Analysis* (Chambers, Cleveland, Kleiner, and Tukey 1983), *The Visual Display of Quantitative Information* (Tufte 1983), and *The Elements of Graphing Data* (Cleveland 1985). We have seen sequels and second editions from Cleveland and Tufte and other works in the same vein (e.g., those by Howard Wainer), but arguably little that has had, or rather deserves, quite so much impact as the four key works published around 1980. The main exception is the book under review, the second edition of one first published in 1999.

Plaudits first: Leland Wilkinson's *The Grammar of Graphics* is one of the richest books available on statistical graphics. It is stuffed full of ideas, information and examples, and any scientist or statistician with a concern for graphics should find something new to them of real use or interest. The title page rightly boasts of 410 illustrations, 319 in full colour, and as a whole the book is very nicely produced. *Grammar* is Wilkinson's *magnum opus*, the fruit of thirty-something years' work in the field, reflecting most of a career's worth of reading and thinking, programming and practice. No other book so nearly matches the multiple aspects of the field.

I say Wilkinson, but should stress that Graham Wills, Dan Rope, Andrew Norton and Roger Dubbs are listed as co-authors on the title page, but not the cover. For those who don't know of Leland Wilkinson, his CV at `http://www.spss.com/research/wilkinson/Publications/wilkinsonvita.pdf` and a detail on p. 98 give the basic details. He was born in 1944, gained degrees at Harvard and Yale, taught Psychology at the University of Illinois at Chicago between 1976 and 1991, and was President of SYSTAT Inc. from 1984 to 1994 until it was bought by SPSS, on which he joined the latter as a Vice-President. Knowing this as background will help readers better understand the book. In essence, Wilkinson has been in and out of academia, and it shows. Positively, half-a-career's worth of teaching and thought is reflected here. However, this book does not grow out of courses and I doubt that it is suitable as a text for teaching, except as wider reading. Naturally that does not rule out it being interesting and useful to a large variety of professionals and researchers.

The reference list is particularly impressive. I count 661 references (although note on the side several omissions from the endlist and several small slips over dates). Naturally, many readers will find a few surprising omissions. I am always puzzled by the frequency of comments (here on p. 194) to the effect that the importance of aspect ratio is a relatively recent discovery. It was clearly emphasised by Fisher (1925) in *Statistical Methods for Research Workers*, and no doubt it was old hat even then.

However, *Grammar* is a flawed masterpiece. Most readers will find it too long by far, although they will naturally differ on what should be left out. That said, it may seem perverse to report one sense in which it is too short. For once, I really missed that curse of many documents, an executive summary of key points. Wilkinson is a very clear writer who wants to explain and entertain, but despite his best efforts a few examples remain unreadable or obscure (e.g., pp. 446, 476, 503).

Similarly, Springer have seemingly given Wilkinson free rein to write the book he wanted, but the result is aimed at a mix of quite different audiences and is presented in a range of styles. Correspondingly, several sections are likely to appear too elementary or too enigmatic to many readers. Is the reader who needs definitions of intersection and union of sets on p. 26 the same reader who will want to check a page and a half of algebraic derivations on pp. 69–70? Is the reader who needs a grounding in basic probability on p. 453 the same reader who will appreciate terse explanations of the EM algorithm and multiple imputation on p. 465? My guess is no in both cases.

Throughout many chapters graphs are described using a formal notation. Only in Chapter 18 does it become completely clear that this is the GPL mentioned briefly in the Preface. A search of http://www.spss.com/ on 2006-12-15 revealed that the company added support for GPL in SPSS 14.0, but they seem unable to decide whether it means Graphics Procedure Language, Graphics Production Language or Graphics Programming Language. Perhaps their marketing people should talk more to the developers. More seriously, I guess that Wilkinson, even as a company Vice-President, had to keep quiet on intentions for GPL until an announcement had been made. No doubt GPL deserves a detailed, separate review by someone who has used it: fairly or not, I will say no more about a language I have never applied.

Wilkinson returns intermittently in this book to a general theme that there is an over-arching grammar of graphics, but quite what he means by this remains elusive as far as I am concerned. I would rather applaud his repeated insistence that in graphics it is the details that matter. Even his organisation into chapters often seems arbitrary. In a strong sense, this book is best seen as a collection of several dozen short pieces.

One reason this review is delayed is that twice in my reading I got bogged down somewhere in the middle, and then felt obliged to start all over again. Only on the third reading did I get all the way through. *Grammar* may be a great novel, in which case I have to report that I was not gripped by the flow and did not understand it as such. I am happy to recommend it as a most rewarding collection of short stories. The best way to see this book is to grasp that Wilkinson is a kind of intellectual magpie, an inveterate collector of material from reading and from his own experiments. The upside of that comes from unexpected encounters and excursions in the middle of something quite different, and the downside from a polymathic perversity, a parade of learning that can be distracting or irritating. (Worryingly, as a kind of magpie myself, I doubt that this review avoids a half-conscious echo of some of these habits.)

Predictably, many readers will tackle this book by skimming to find graphs that look worthwhile for their own purposes, and then reading the associated text. Broadly, this is the best way for most readers to approach the book, however philistine the advice may sound. Naturally, it doesn't apply to most classics, in statistics or elsewhere. Skimming *War and Peace* would not convey more than that many things happen to several Russians over a long period, but skimming *The Grammar of Graphics* will be much better than ignoring it because you

lack the time to read it thoroughly.

Who is this book for? More should be said about that. Often with technical books, it is obvious or it does not matter much. A great variety of people could get *something* out of all the other books mentioned so far. Here it is not obvious and it does matter. On pp. x–xi first one general audience is identified as 'anyone who is interested in business or scientific graphics' and then three specific audiences are identified, as college and graduate students in computer sciences and statistics ('This is the only book in print that lays out in detail how to write computer programs for business or scientific graphics.'); mathematicians, statisticians and computer scientists who are not experts in quantitative graphics; and statistics and computer science specialists in graphics. Having read the book, I can say that all these audiences will indeed find much material that is aimed in their direction, and similarly much that is not. Least plausible is the claim that students will learn from *Grammar* how to write graphics programs. For that they would need not only to know a suitable language, which would have to be learned elsewhere, but more crucially to have enough scientific and statistical maturity to design such a system. Besides, how many students would be well-advised to start writing new programs, as compared with learning how to write code in some pre-existing graphical language or system? I fear I am missing Wilkinson's point here.

Wilkinson returns to the question of what he is trying to do on pp. 13–18: the book is not a command language, not a taxonomy, not a drafting package, not a book of virtues, not a heuristic system, not a geographic information system, not a visualization system. That is a lot of 'not's for a book nearly 700 pages long and adds to the impression of complexity and idiosyncrasy: the author is trying to do several different things, yet also definitely not trying to do several other different things.

The negative here that to me bites hardest is the denial that this is a book of virtues. Most of the leading authors on statistical graphics are very confident that there are good graphs and bad graphs and very focused on conveying their views on the differences to readers. The methods for doing so can vary – subtle or sensitive charm, scathing criticism or simply ignoring graph types thought to be poor – but the aims are similar, even if the authors do not all agree! For reasons I do not fully grasp, Wilkinson differs. Indeed, at times, he even ignores obvious simple and direct displays in favour of experimenting with unusual alternatives, seemingly the more unusual the better.

This trait can be seen in several examples, particularly when data sets are revisited intermittently throughout the book. Three categorical data sets (ACLS data on pp. 31, 35–38, 172, 208, 211–212; sex partners and attitudes to Bible on pp. 175, 296, 308, 474, 478; Titanic data on pp. 342–343, 403–405) are plotted in all sorts of different ways, but many of them strike me as unsuccessful, even perverse. The basic but still useful idea of a table of bars hardly gets a look in.

But Wilkinson also has a side that treasures simplicity and directness. One of his best known papers is on dot plots, which in my view usually work better than box plots for one or a few groups of data. Hence it is a pleasure to see nice straightforward examples: a dot plot of correlations (p. 45), the advice to start with dot plots (p. 137) and a composite dot-box plot (p. 307). Similarly, comparing male and female crime data by a difference of fractions (p. 58), namely fraction of crimes committed by males minus fraction of crimes committed by females, is a good pattern to follow.

A particular strength of this book is attention to data spaces that are neglected by several

key texts in statistical graphics, including multivariate, spatial, temporal, triaxial (barycentric coordinates) and circular. Wilkinson's background and own research underpins the first of these in particular, which is richly represented.

As a professional geographer, I especially applaud Wilkinson's generous recognition of cartographic ideas and examples. Predictably, I have a few quibbles. The so-called Peters projection does not favour the Southern Hemisphere (p. 241), as it treats both hemispheres similarly. Choropleth (patch) maps should show density, not absolute amounts (pp. 315–6): otherwise size of area is compounded confusingly with intensity of shading.

Circular scales, such as compass direction or time of year, partly define some of these data spaces. Thus data on relative humidity as a function of wind direction are shown using conformal mapping (p. 204). However, the graphic is not evaluated critically. Although evidently an attempt to match the data space, it is arguably less helpful in practice than a standard scatter and line plot would be, for all that the latter would cut the circle quite arbitrarily. In this particular example, few winds blow from a sector of about 90° centred just East of North, so the cut would not be painful. The broader claim that conformal mapping can be very helpful, even 'addictive' (p. 202), falls flat without a range of statistical examples. Similarly, a polar plot of time series (p. 214) is not really successful. As the series shows a strong trend and no obvious seasonality, either the example or the format is poorly chosen. Finally on this point, the example on p. 341 shows polar plots of wind direction themselves arranged in a polar array by month of the year. The idea is ingenious, and indeed droll, but also ineffective: so little space is devoted to showing the data that only the grossest contrasts can be decoded easily.

I noticed a few statistical slips. I don't understand how a gamma distribution can be defined by a single parameter (p. 109). The discussion of estimation of the standard deviation needs tidying up on p. 460 to correct the implication that the usual recipe based on the sample size minus 1 is unbiased. The variance formula on p. 464 is incorrect.

These slips are small beer compared with the statistical pluses in this book, including some excellent turns of phrase or little stories. Consider this trumpet-blast for statistical thinking on p. 354:

'It is not always convenient to remember that the right model for a population can fit a sample of data *worse* than a wrong model—even a wrong model with fewer parameters. We cannot rely on statistical diagnostics to save us, especially with small samples. We must think about what our models mean, regardless of fit, or we will promulgate nonsense.'

Or consider an intriguing list of definitions of twelve words indicating different facets of uncertainty (p. 452). However, I doubt that few statistical scientists would accept as part of a definition of error that it 'varies to left or right of the truth with equal likelihood', unless this is a little political joke. Nor is accuracy for all disciplines defined as 'relative lack of bias and error'.

Other gems include nice stories on how FBI statisticians uncovered electoral fraud in Chicago (not using anything fancy, but find out how yourself) (p. 57) and the reports of 88 acts of coitus that may well be keypunch errors for a missing data code of 99 (p. 475). The discussion on scale breaks on pp. 351–2 is also particularly good, although not I think quite everything that could be said.

Terminology can often be a problem, distracting attention from more vital issues. You may not agree that the term 'quantile plot' applies to plots of cumulative distribution functions

(pp. 91–92, 587). The distinction between parallel coordinates plots (p. 251) and profile plots (p. 314) likewise is a little difficult to swallow.

Given the length and complexity of this work, it is not surprising to find a few examples of contradiction and repetition. The latter, of course, may help some readers. Graphical lying about the data is deplored on p. 6 and practised on p. 12. Double y-axes are similarly deplored on p. 352 but in evidence on pp. 432–3. Andrews curves are covered twice over, on pp. 250 and 302. The horseshoe phenomenon is also covered twice, on pp. 250 and 529.

Given the excellent general standard of presentation, it is correspondingly a shame that the text was not more carefully copy-edited.

Despite a section about units (pp. 86–89), Wilkinson is a little careless in practice about labelling graphs with units. Sometimes units are not specified at all. The abbreviation 'gm' is often used for grammes: this is long since non-standard ('g' should be used). On some graphs Fahrenheit temperatures are not specified as such (pp. 140–142, 299, 556, 613). Most astonishingly, bar heights are stated to be proportional to temperatures (pp. 303–304), which to say the least depends on use of a Fahrenheit scale and zero and negative temperatures not being observed.

Wilkinson's learning and curiosity are shown by etymological explanations of the Greek, Latin and other roots of keywords from most chapter titles at the beginnings of chapters. One is muddled: In ancient Greek, as in English, 'analysis' is a noun, not a verb (p. 489). This enthusiasm for words has not stopped some extraordinary word choices: 'eponomy' for 'eponymy' (p. xiv), 'diladic' for 'dyadic' (p. 29), 'dilatation' for 'dilation' (p. 191: cf. p. 202), 'nemesis' for 'antagonist' (pp. 233, 270), 'rectangular' for 'oblong' (pp. 377, 497). Linguistic conservatives will shudder at the frequency of 'comprised of' and 'compared to'. Accents should be restored to several surnames (Cézanne, p. 6; Šalkauskas, pp. 118, 160; Sierpiński, pp. 374, 391; Poincaré, pp. 422, 559). There are many other minor typos, such as translating Wheatstone a century forward to 1938 (p. 232).

An international publisher should temper the tendency of some authors to assume that readers are based in the United States. What I guess are little jokes about Pittsburgh (p. 27), Philadelphia (p. 183) and Boston (p. 234) all escaped me, despite a training as a geographer and several trips across the Atlantic. Other examples of assuming United States perspectives occur on pp. 53 and 355. The last mention of wit in the review should, however, be positive. Wilkinson has a nice sense of humour that adds pleasure to reading his book. It will be no surprise that not everybody sees every joke.

The author index shows some split personalities. The subject index is poor, and many readers will find themselves adding extra entries to aid their return visits. However, note the useful entry on data examples on p. 682.

The Grammar of Graphics is a long, complex and contradictory book. As you can see, I have been unable to avoid writing a matching review. Lest this evaluation appear dominated by a mass of picky or pedantic criticisms, let me close with applause. Warts and all, *The Grammar of Graphics* is a richly rewarding work, an outstanding achievement by one of the leaders of statistical graphics. Seek it out. (And let us see whether the publisher quotes 'Warts and all'.)

References

Chambers JM, Cleveland WS, Kleiner B, Tukey PA (1983). *Graphical Methods for Data Analysis*. Wadsworth.

Cleveland WS (1985). *The Elements of Graphing Data*. Wadsworth.

Fisher RA (1925). *Statistical Methods for Research Workers*. Oliver and Boyd.

Tufte ER (1983). *The Visual Display of Quantitative Information*. Graphics Press.

Tukey JW (1977). *Exploratory Data Analysis*. Addison-Wesley.

Reviewer:

Nicholas J. Cox
Department of Geography
Durham University, United Kingdom
E-mail: n.j.cox@durham.ac.uk

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 17, Book Review 3 *Published:* 2007-01-15
January 2007