

Edward R. Tufte

The Visual Display of Quantitative Information

SECOND EDITION

Graphics Press · Cheshire, Connecticut

Copyright © 2001 by Edward Rolf Tufte
PUBLISHED BY GRAPHICS PRESS LLC
POST OFFICE BOX 430, CHESHIRE, CONNECTICUT 06410

All rights to illustrations and text reserved by Edward Rolf Tufte. This work may not be copied, reproduced, or translated in whole or in part without written permission of the publisher, except for brief excerpts in connection with reviews or scholarly analysis. Use with *any* form of information storage and retrieval, electronic adaptation or whatever, computer software, or by similar or dissimilar methods now known or developed in the future is also strictly forbidden without written permission of the publisher. A number of illustrations are reproduced by permission; those copyright-holders are credited on page 197.

Contents

PART I GRAPHICAL PRACTICE

- 1 *Graphical Excellence* 13
- 2 *Graphical Integrity* 53
- 3 *Sources of Graphical Integrity and Sophistication* 79

PART II THEORY OF DATA GRAPHICS

- 4 *Data-Ink and Graphical Redesign* 91
- 5 *Chartjunk: Vibrations, Grids, and Ducks* 107
- 6 *Data-Ink Maximization and Graphical Design* 123
- 7 *Multifunctioning Graphical Elements* 139
- 8 *Data Density and Small Multiples* 161
- 9 *Aesthetics and Technique in Data Graphical Design* 177
- Epilogue: Designs for the Display of Information* 191

For my parents

Edward E. Tufte and Virginia James Tufte

To the memory of

John W. Tukey (1915–2000)

Introduction to the Second Edition

This new edition provides high-resolution color reproductions of the many graphics of William Playfair, adds color to other images where appropriate, and includes all the changes and corrections accumulated during the 17 printings of the first edition.

This book began in 1975 when Dean Donald Stokes of Princeton's Woodrow Wilson School asked me to teach statistics to a dozen journalists who were visiting that year to learn some economics. I annotated a collection of readings, with a long section on statistical graphics. The literature here was thin, too often grimly devoted to explaining use of the ruling pen and to promulgating "graphic standards" indifferent to the nature of visual evidence and quantitative reasoning. Soon I wrote up some ideas. Then John Tukey, the phenomenal Princeton statistician, suggested that we give a series of joint seminars. Since the mid-1960s, Tukey had opened up the field, as his brilliant technical contributions made it clear that the study of statistical graphics was intellectually respectable and not just about pie charts and ruling pens.

After moving to Yale University, I finished the manuscript in 1982. A publisher was interested but planned to print only 2,000 copies and to charge a very high price, contrary to my hopes for a wide readership. I also sought to design the book so as to make it *self-exemplifying*—that is, the physical object itself would reflect the intellectual principles advanced in the book. Publishers seemed appalled at the prospect that an author might govern design.

Consequently I investigated self-publishing. This required a first-rate book designer, a lot of money (at least for a young professor), and a large garage. I found Howard Gralla who had designed many museum catalogs with great care and craft. He was willing to work closely with this difficult author who was filled with all sorts of opinions about design and typography. We spent the summer in

his studio laying out the book, page by page. We were able to integrate graphics right into the text, sometimes into the middle of a sentence, eliminating the usual separation of text and image—one of the ideas *Visual Display* advocated. To finance the book I took out another mortgage on my home. The bank officer said this was the second most unusual loan that she had ever made; first place belonged to a loan to a circus to buy an elephant!

My view on self-publishing was to go all out, to make the best and most elegant and wonderful book possible, without compromise. Otherwise, why do it?

Most of all, the book, as a thing in itself, gave to me fresh new eyes for the intellectual and aesthetic joy of visual evidence, visual reasoning, and visual understanding.

January 2001

Cheshire, Connecticut

Introduction

Data graphics visually display measured quantities by means of the combined use of points, lines, a coordinate system, numbers, symbols, words, shading, and color.

The use of abstract, non-representational pictures to show numbers is a surprisingly recent invention, perhaps because of the diversity of skills required—the visual-artistic, empirical-statistical, and mathematical. It was not until 1750–1800 that statistical graphics—length and area to show quantity, time-series, scatterplots, and multivariate displays—were invented, long after such triumphs of mathematical ingenuity as logarithms, Cartesian coordinates, the calculus, and the basics of probability theory. The remarkable William Playfair (1759–1823) developed or improved upon nearly all the fundamental graphical designs, seeking to replace conventional tables of numbers with the systematic visual representations of his “linear arithmetic.”

Modern data graphics can do much more than simply substitute for small statistical tables. At their best, graphics are instruments for reasoning about quantitative information. Often the most effective way to describe, explore, and summarize a set of numbers—even a very large set—is to look at pictures of those numbers. Furthermore, of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful.

The first part of this book reviews the graphical practice of the two centuries since Playfair. The reader will, I hope, rejoice in the graphical glories shown in Chapter 1 and then condemn the lapses and lost opportunities exhibited in Chapter 2. Chapter 3, on graphical integrity and sophistication, seeks to account for these differences in quality of graphical design.

The second part of the book provides a language for discussing graphics and a practical theory of data graphics. Applying to most visual displays of quantitative information, the theory leads to changes and improvements in design, suggests why some graphics might be better than others, and generates new types of graphics. The emphasis is on maximizing principles, empirical measures of graphical performance, and the sequential improvement of graphics through revision and editing. Insights into graphical design are to be gained, I believe, from theories of what makes for excellence in art, architecture, and prose.

This is a book about the design of statistical graphics and, as such, it is concerned both with design and with statistics. But it is also about how to communicate information through the simultaneous presentation of words, numbers, and pictures. The design of statistical graphics is a universal matter—like mathematics—and is not tied to the unique features of a particular language. The descriptive concepts (a vocabulary for graphics) and the principles advanced apply to most designs. I have at times provided evidence about the scope of these ideas, by showing how frequently a principle applies to (a random sample of) news and scientific graphics.

Each year, the world over, somewhere between 900 billion (9×10^{11}) and 2 trillion (2×10^{12}) images of statistical graphics are printed. The principles of this book apply to most of those graphics. Some of the suggested changes are small, but others are substantial, with consequences for hundreds of billions of printed pages.

But I hope also that the book has consequences for the viewers and makers of those images—that they will never view or create statistical graphics the same way again. That is in part because we are about to see, collected here, so many wonderful drawings, those of Playfair, of Minard, of Marey, and, nowadays, of the computer.

Most of all, then, this book is a celebration of data graphics.

PART I

Graphical Practice