BOOK REVIEW

If a picture is worth a thousand words, how many should a chart be worth? A good one can summarise thousands of data points and highlight a causal relationship; poor ones can obscure a few simple numbers with meaningless "chartjunk", adding nothing but confusion and irritation. **The Visual Display of Quantitative Information**, one of Amazon's books of the century, is essential reading for designers, analysts, and managers who want to produce better graphs.

Rescuing data graphics from the artists

When most people think of graphing statistics, if they think of it at all, what comes to mind is probably How to Lie With Statistics. But the problem is far deeper than deliberate deceit. The use of data graphics as a gimmick to hold the attention of readers assumed to be too dull-witted to understand numbers has led to a triumph of style (or fluff) over substance. Too many graphics are either patronisingly simplistic, drowning in decoration or downright misleading. Edward Tufte is on a mission to set that right, starting with a couple of rules:

- If the statistics are boring then you've got the wrong numbers
- You should respect the intelligence of your audience

Most bad graphics don't set out to mislead. For example any graph purporting to show a change in price that charts actual as opposed to "real" cost is always misleading, even though it reflects the original numbers perfectly well. Accurate graphing takes thought and experience as well as integrity, which is why it should be a task for those who can combine an aesthetic touch with some statistical and analytical expertise.
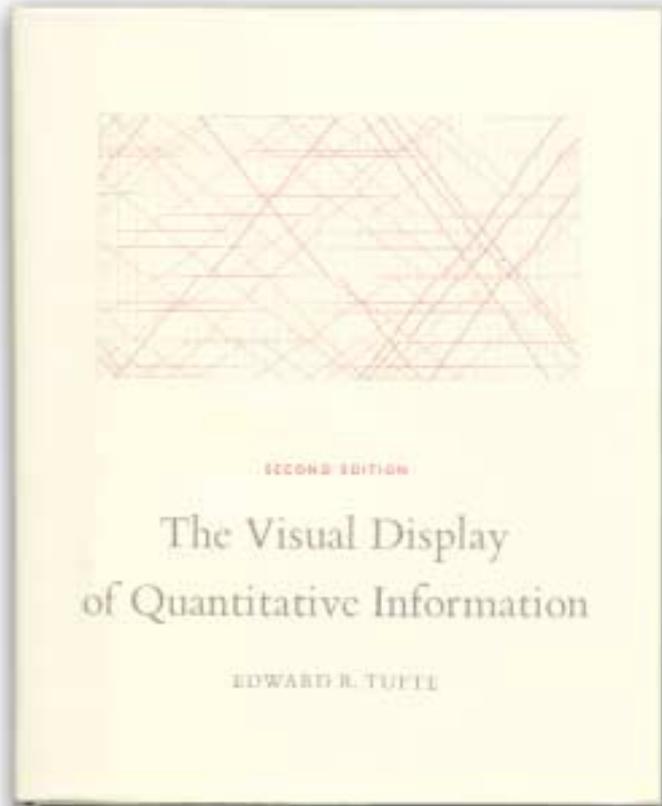
Tufte has a knack for firing the enthusiasm of those who read his work. His style is witty and engaging, occasionally becoming acerbic when discussing a particularly bad example. Of one truly monstrous 3D graphic from the pages of American Education he comments: *"five colors report, almost by happenstance, only five pieces of data....This may well be the worst graphic ever to find its way into print."*

It certainly is hideous, but dating from the 1970s it seems in retrospect only a warning of what was to come with the advent of desktop computing and Chart Wizards.
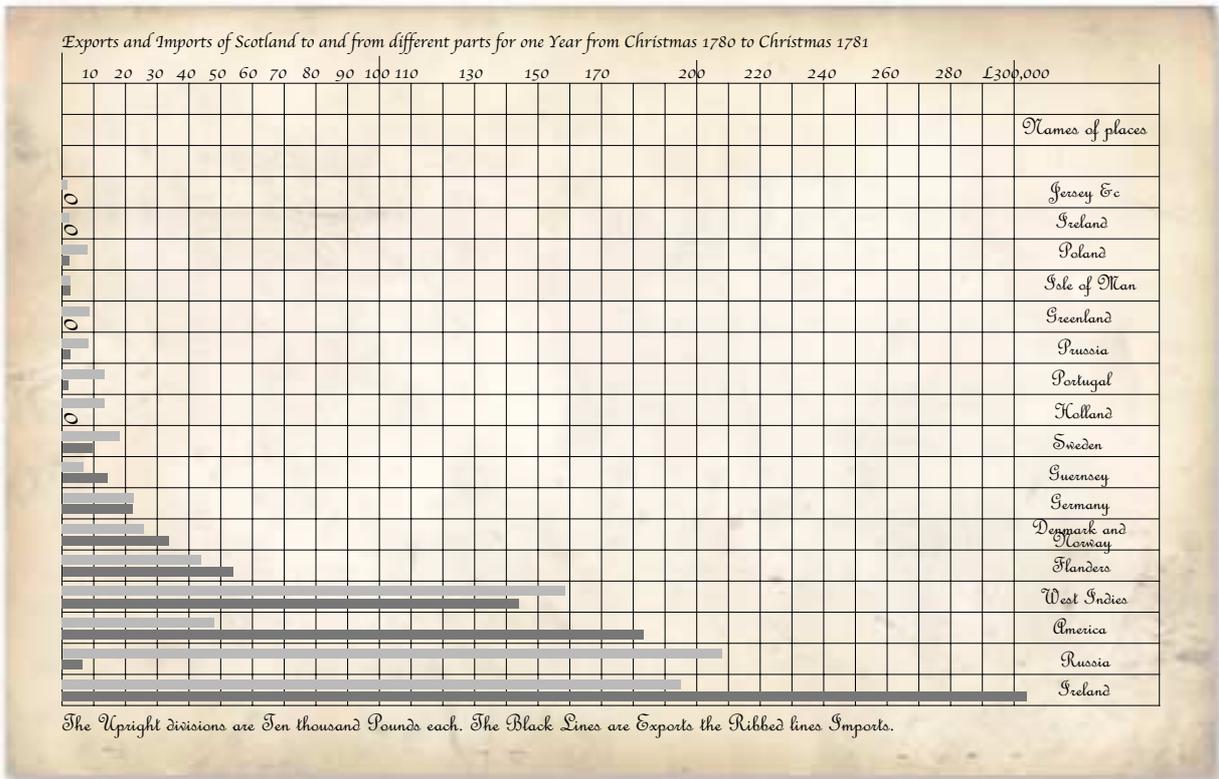
The good, the bad and the ugly

Tufte raids his considerable library of rare books to illustrate good and bad design. The reminder that ideas such as the bar chart, time series and scatter plot were inventions rather than solutions handed

Playfair Bar chart

Exports and Imports of Scotland to and from different parts for one Year from Christmas 1780 to Christmas 1781

| 10 20 30 40 50 60 70 80 90 100 110 | 130 | 150 | 170 | 200 | 220 | 240 | 260 | 280 £300,000 |

Names of places

Jersey &c
Ireland
Poland
Isle of Man
Greenland
Prussia
Portugal
Holland
Sweden
Guernsey
Germany
Denmark and Norway
Flanders
West Indies
America
Russia
Ireland

The Upright divisions are Ten thousand Pounds each. The Black Lines are Exports the Ribbed lines Imports.

Tufte on PowerPoint

Tufte hates PowerPoint. Not the software itself, but its use for unsuitable material, the presentation style it encourages and the patterns of thought it cements. To quote a recent essay:

"At a minimum, a presentation format should do no harm. Yet the PowerPoint style routinely disrupts, dominates, and trivializes content. Thus PowerPoint presentations too often resemble a school play - very loud, very slow, and very simple."

He expands the point in a pamphlet **The Cognitive Style of PowerPoint**, an excerpt from his forthcoming **Beautiful Evidence**. In brief he feels that PowerPoint is a sales tool, incapable of sufficient resolution to present meaningful statistical data and forcing a linear style that hampers our ability to see patterns and relationships. In addition the bullet-point style tends to trivialise and simplify rather than illuminate.

It is revealing, and terrifying, that both Tufte and Richard Feynman saw a direct link between NASA's use of bullet-style thinking and the two shuttle disasters. The Columbia Accident Investigation Board agreed - PowerPoint is inadequate for technical communication.

down from on high is a salutary one that should prompt us to ask if things can be made better.

The example above may be the first bar chart? Playfair was unconvinced, but the idea has stood the test of time

The choice of graph really does matter. Before the scatter plot how was it possible to examine the relationship between two variables? Fundamentally it wasn't. The use of the right chart could make a significant difference to your analysis, particularly with complex multivariate data.

Tufte provides examples of good practice ranging from Galileo to 3D computer

BOOK REVIEW
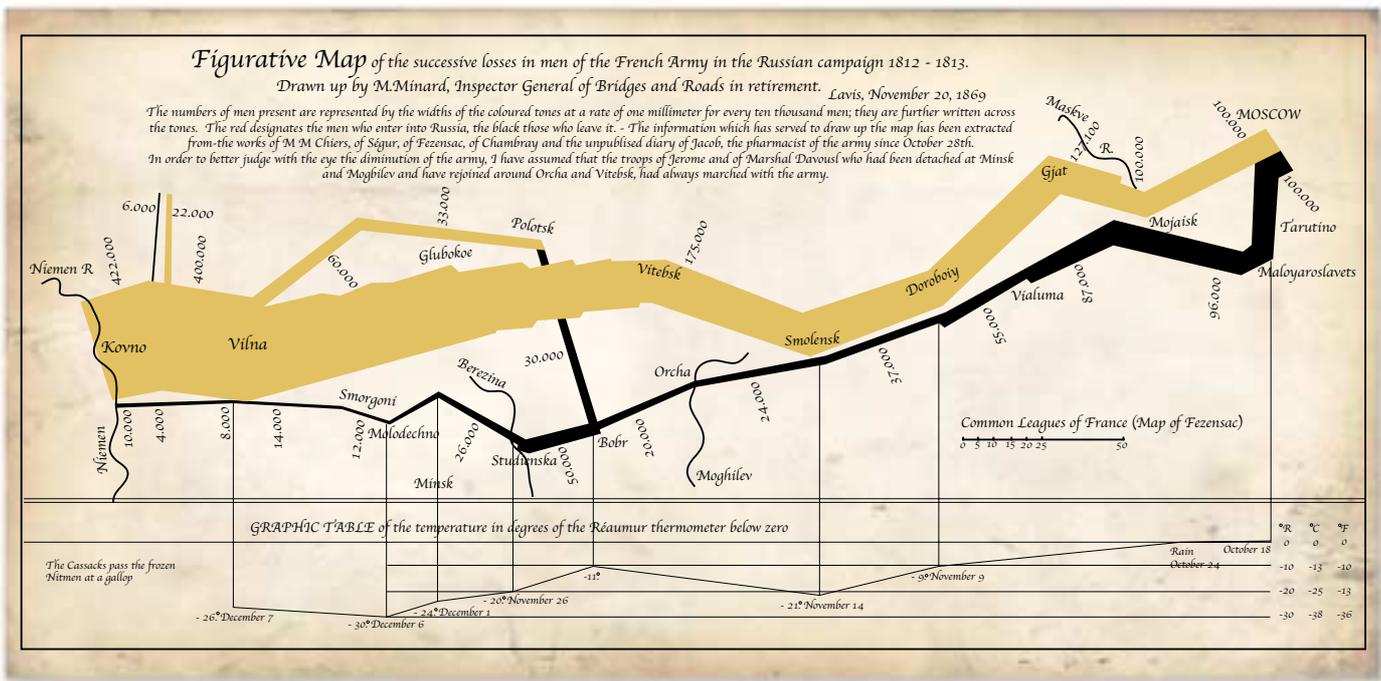
Minard: The 1812 campaign (translated)

visualisations, some of which are truly inspirational. It is revealing that the very best data graphics are often one-offs; works of art that bring to life a specific set of data. Minard's 1869 map above narrates Napoleon's disastrous 1812 campaign:

Combining chart, map and time-series it chronicles the terrible losses of the French army in the diminishing width of the line showing advance and retreat. According to Tufte it "may well be the best statistical graphic ever drawn." It certainly has a stark power that would be difficult to match in words.

Tufte does not just lay down principles and show examples of good and bad practice. Perhaps the most useful sections of the book are those in which he

elaborates some of his ideas for redesigning standard graphs step by step. This takes the work from the realm of theory to practical application, and some of his suggestions are enormously valuable, deserving wide adoption. His latest innovation, the sparkline (to be featured in the forthcoming **Beautiful Evidence**) is already gathering considerable momentum, and may be the first of his suggestions to reach the mainstream.

His main aim in making changes is to maximise the "data-ink ratio". Any mark on the paper should add to the information contained, if it does not then it can be erased. A good example is the overbearing grid that dominates so many charts. Sometimes a light grid may help when looking up specific data, but a

Sparklines

Tufte describes sparklines as "data-intense, design-simple, word-sized graphics". The neat part is that they can fit into a flow of text naturally, appearing at the appropriate point to illustrate an argument. They contain nothing but data - no axes and little labelling, but they still manage to tell a clear and rich story.

This one ～～ CSI 84.2, for example, shows the trend of customer satisfaction over the last 15 months. With the addition of a grey stripe to indicate the normal range it also becomes a tool to spot "out of bounds" values . This is especially useful when glancing at a range of figures simultaneously. What could be more appropriate for a strategic "dashboard" that is high on information and clarity but low on gimmicks?

～～	Customer satisfaction	84
～～	Employee satisfaction	76
～～	Brand evaluation	79
～～	Share price	320

Biography

As a professor in the fields of statistics and graphic design, Tufte has reinvented a field by bringing the two together. For too long graphics have been the exclusive preserve of artists with no understanding of statistics. Tufte's aim is too foster a discipline that can combine the talents of the artist and the statistician.

Tufte is the worlds leading expert on the effective visual presentation of statistical information, and his books have become instant classics. A hallmark of his approach is that his books are self published because he wasn't satisfied that anyone else would take the necessary care in laying out and printing them. The results are works of art in themselves, elegantly typeset and well bound.

Bibliography

Visual Explanations
Envisioning Information
The Visual Display of Quantitative Information
Beautiful Evidence. (forthcoming)
http://www.edwardtufte.com

A good measure of graphical excellence is data density, almost precisely a measure of how many words the chart is worth. If the data density is lower than a table, then the chances are a table would be better.

Does it really matter?

Perhaps you're thinking that this is a diatribe that we can afford to dismiss. The way data is presented surely can't make all that much difference? I would argue that it is hard to overestimate the importance of this area. The point is that bad graphics can not only be made to lie, but frequently mislead or hide information inadvertently because of poor design, and this has daily consequences for the decisions we make.

Is the way you present data really important? Yes, if you're talking about data on cancer death rates. Yes, if you're talking about the key information relating to the safety of a NASA shuttle mission. Yes, if you're talking about any data that matters at all. And if doesn't, then why are you graphing it in the first place? **S**

heavy grid just gets in the way.

Rules for effective graphing

Tufte's approach boils down to one simple rule - graphs should free up numbers to speak for themselves, not obfuscate or decorate for the sake of it. Some key principles:

Above all else show the data
Maximise the data-ink ratio
Erase non-data-ink
Erase redundant data-ink
Revise and edit

Charts that explain

Tufte points out that the best charts are those which reveal something that could not easily be seen in the data alone. Simple data are often best shown in the most simple way - a table. Pie charts, by this rationale, are rarely a good use of ink. Of course tables present a typographic challenge that is worthy of a book in itself, and are often just as poorly produced as the average graph!

The point is that numbers can best be made interesting not by turning them into

asinine pictures, but by facilitating comparisons and highlighting relationships. Simple tables and pie charts don't do this well, but intelligently designed tables, time series and multivariate data plots are all capable of revealing new and interesting truths.

Before and after -
Tufte redesigns a population graph