
The background of the slide is a black field filled with a repeating pattern of white, stylized geometric shapes. These shapes are complex, multi-pointed star-like figures or fractal-like patterns, each composed of many small triangles and lines. They are arranged in a grid-like fashion, creating a dense, textured background.

Edward Tufte

The Visual Display of Quantitative Data

Second Edition, Ninth Printing, July 2015

By Adrian Crockett, November 2018

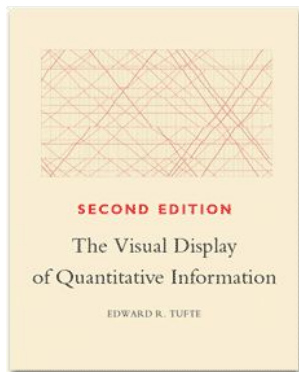
Who is Edward Tufte?

*An American
statistician and
Professor Emeritus at
Yale University
(political science,
statistics, and computer
science) ...and in his
spare time an artist*



- **Born:** 14th March 1942 in Kansas City, Missouri
 - **Education:**
 - BS Statistics (Stanford University)
 - MS Statistics (Stanford University)
 - PhD Political Science (Yale University)
 - **Teaching:**
 - Princeton University: Political Economy and Data Analysis
 - Yale University: Political Science, Statistics and Computer Science
-

The course materials and related lectures developed with [John Tukey](#) became the foundations of...



The Tipping Point

- **Journalists:** Tufte was asked in 1975 by Princeton to teach statistics to visiting journalists to increase their knowledge of economics
 - **Ignorance:** This was at a time when statistics were seen as boring, and the perception of most journalists was that they needed to dumb down their materials for readers rather than add rigor
 - **The future:** In his first book discussed today, Tufte strongly refutes this *a priori* assumption Chapter 3
-

Tufte's goals for the book

Given Tufte's contributions to the data visualization field, it's interesting to think what the field would look like today if that bank officer had turned him down...



- **Large circulation:** His purpose wasn't to appeal only to academics
 - **Design specificity:** He needed control over the design of the book - something unconventional at the time
 - **Rejection:** No publishers wanted any part in this
 - **The result:** Self-publication
 - **The commitment:** Second mortgage on his house to finance the book
-

Breaking down the book

The book is broken down into two main parts, the first of which provides some heuristics on graphical practice...

...while the second part provides a theory of data graphics and formalizes the practice...

...and the epilogue offers forgiveness

Graphical
Practice
(Heuristics)

Graphical
Excellence

Graphical Integrity

Sources of
Graphical Integrity
and Sophistication

Theory of Data
Graphics
(Rules)

Data-Ink and
Graphical
Redesign - Rules 1-3

Chartjunk:
Vibrations, Grids,
and Ducks - Rules 4-6

Data-Ink
Maximization and
Graphical Design

Multifunctioning
Graphical
Elements

High-Resolution
Data Graphics
- Rules 7-8

Aesthetics and
technique in Data
Graphical Design -
Rules 9-10

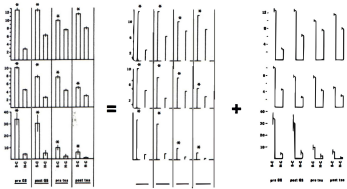
Epilogue
(Forgiveness)

“The principles should not be applied rigidly or in a peevish spirit;...it is better to violate any principle than to place graceless or inelegant marks on paper” p. 191

Data-Ink and Graphical Redesign

Overarching principle:
“Above all else show the data” p.92

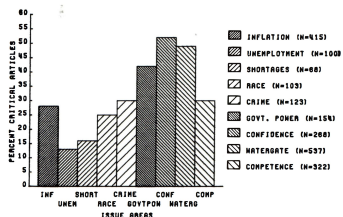
*In the example below
Tufte shows the
subadditivity of the
reduction rules (#2a &
b)...p. 102*



- “**Data-Ink** is the non-erasable core of a graphic, the non-redundant ink arranged in response to variation in the numbers presented” p. 93
 - **Data-Ink Ratio:** Data-ink/total ink used to print the graphic
 - **Rule #1:** Maximize the data-ink ratio, *within reason*
 - **Rule #2:** Two erasing principles
 - **#2a:** Erase non-data ink, *within reason*
 - **#2b:** Erase redundant data-ink, *within reason*
 - **Rule #3:** Revise and edit
-

Chartjunk: Vibrations, Grids, and Ducks

Tufte is generally negative on software packages for the masses and is concerned that they unleash charts like this with unintentional moire vibrations...

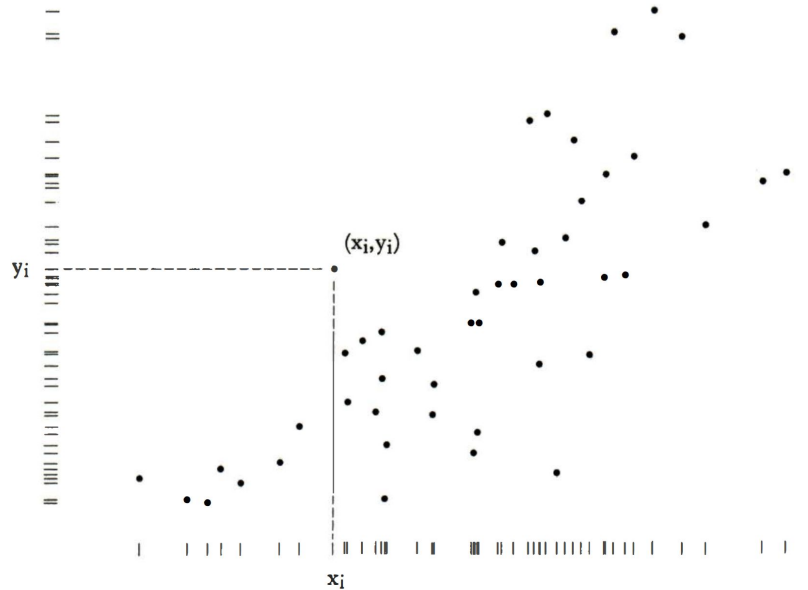


- **Rule #4:** Forgo [moire vibrations](#) (feels dated)
 - In the early ages of computer graphics and grayscale printing, many software packages defaulted to using patterns that unintentionally created moire vibrations
- **Rule #5:** Forgo the grid
 - Wherever possible forgo the grid
 - If it is necessary (for instance [Marey train schedule](#)) lighten and thin the grid lines see Rule #9
- **Rule #6:** Forgo the duck (chartjunk)
 - “When a graphic is taken over by decorative forms or computer debris...then that graphic may be called a [duck](#)” p 116

Data-Ink Maximization and Graphical Design

"Is it possible to do what a theory of graphics is supposed to do? That is, to derive new graphical forms"
p. 125

In this chapter Tufte explores the impact of his earlier rules, and how they create new types of charts. The one I've used in practice most is...



In this chapter Tufte offer redesigns for:

The Box Plot

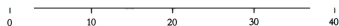
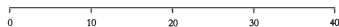
Bar chart/Histogram

Scatterplot

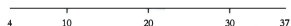
Notice, his focus on “scientific” charts rather than more popular charts from an end users perspective

Multi-functioning Graphical Elements

Overarching principle:
“Mobilize every graphical element, perhaps several times over, to show the data”
p.139



Two different versions of double functioning labels



- Rather than presenting formal rules of a theory, this chapter provides four major examples:
 - Data-Built Data Measures (unusual today)
 - Data-Based Grids (“very occasionally the grid can report data directly”) p. 145
 - **Double-Functioning Labels** (most commonly use - see sidebar)
 - Puzzles and Hierarchy in Graphics (mainly maps)
 - However, even Tufte hedges his bets in this chapter and notes that these are uncommon examples
 - For me this is the least valuable chapter in the book
-

High-Resolution Data Graphics

Overarching principle:
Humans can handle much higher levels of data density than most graphics are generally designed for...

...we should be taking advantage of this, not shying away from it

- **Data Density of Display:** Number of entries in data matrix/area of data display
 - **Rule #7:** For Non-Data Ink, Less is More
 - Maximize data density and the size of the data matrix, *within reason* (but at the same time exploiting the maximum resolution of the available data-display technology)
 - **Rule #8:** For Data-Ink, Less is Bore
 - [Small multiples](#)
 - [Sparklines](#)
-

Aesthetic and Technique in Data Graphical Design

This chapter is a little mixed with some general guidance and rules mixed in...

“Graphical elegance is often found in simplicity of design and complexity of data”
p. 177

- The Choice of Design
 - Sentences, Text-Tables, Tables, Semi-Graphics, Graphics
 - Making Complexity Accessible
 - Combining Words, Numbers and Pictures - “Data graphics are paragraphs about data and should be treated as such” p.181
 - Accessible Complexity - the Friendly Data Graphic
 - Proportion and Scale:
 - **Rule #9:** Line weight and lettering: Use thin by default, and use thick for contrast in meaning
 - **Rule #10:** Shape: move toward horizontal graphics about 50 percent wider than taller, unless the nature of the data suggests the shape of the graphic, in which case follow that suggestion
-

Discussion Questions

While Tufte has been subject to critique over the years, I find most of it unfounded or exaggerated and his contributions to the field speak for themselves

And importantly, don't forget to take his comments in the Epilogue to heart - all the rules are there to guide, as opposed to be strictly adhered to

- What is the style/use/purpose of your subject?
 - Data visualization critic (the most prominent and systematic one)
 - What does this subject do well? What do they do poorly? What do you like/dislike?
 - **Well:** Breaks the problem into easy to digest rules
 - **Poorly:** Doesn't connect with audience (often seems aloof)
 - **Like:** Brought a system where there was none
 - **Dislike:** Materials feel outdated....missing the connection to today
 - How does this subject connect to other readings and discussions within the course?
 - Tufte's work has influenced virtually everything written on data visualization in the last 20 years - it's everywhere
-