

# CS 519: Scientific Visualization

---

## Information Visualization: Design

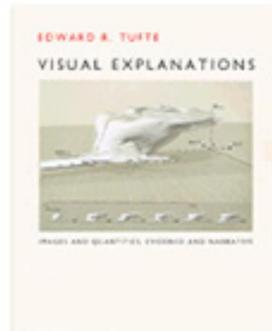
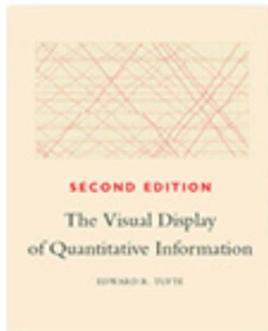
Eric Shaffer

**Slides adapted from:**

- Alvitta Otley, Washington University in St Louis
- Mariah Mayer, University of Utah
- Remco Chang, Tufts University

# Edward Tufte

- Well-known information visualization researcher/practitioner
  - Probably a big understatement
- Evangelist for good design
- Focusses mostly on static visualizations
  - But his principles can often be applied to interactive visualizations



# Principles of Graphical Excellence

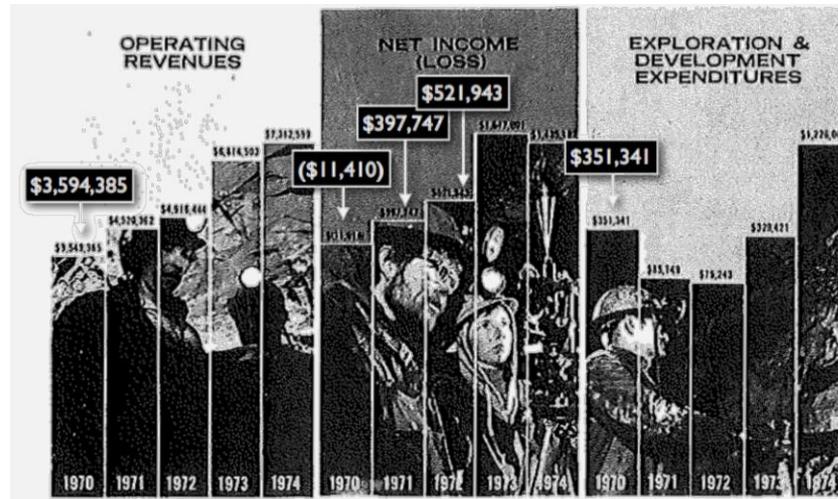
- ▣ Show the data
- ▣ Avoid distorting the data
- ▣ Present many numbers in a small space
- ▣ Make large data sets coherent
- ▣ Encourage the eye to compare different pieces of data
- ▣ Serve a purpose
  - ▣ Description
  - ▣ Exploration
  - ▣ Tabulation
  - ▣ Decoration

From ***The Visual Display of Quantitative Information*** –  
Edward Tufte

# Graphical Integrity

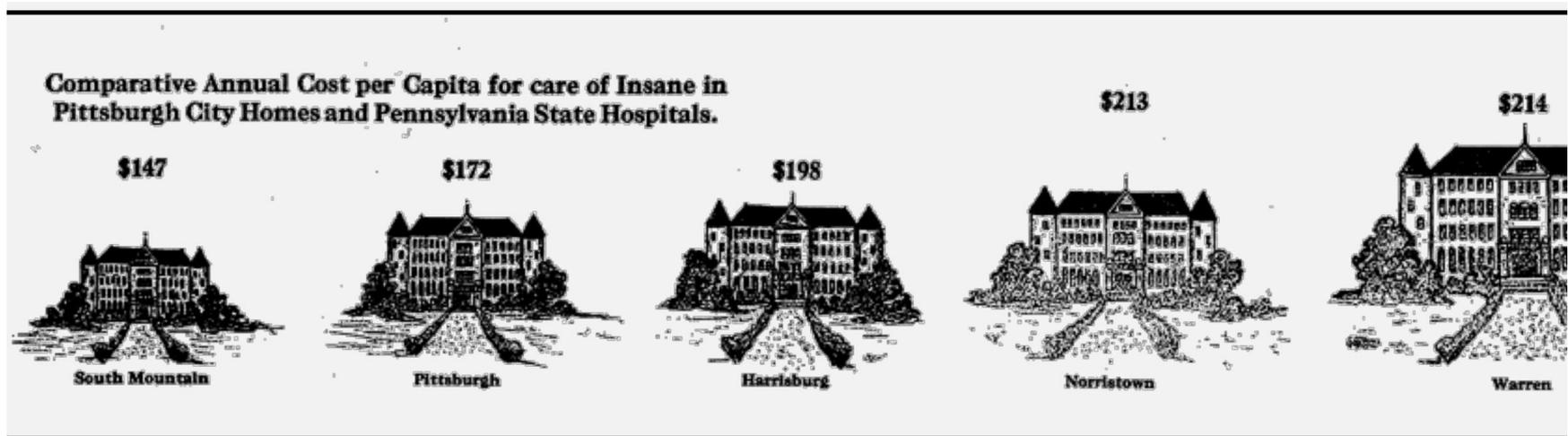
- ▣ Tufte lays out guiding principles...
  - ▣ Clear and detailed labelling should be used to defeat graphical distortion and ambiguity.
  - ▣ “Above all else show the data”
- ▣ So...how can you lie with info visualization?

# Missing Labels



- “...annual report of a company that would just as soon forget about 1970. A careful look at the middle panel reveals a negative income, which is disguised by having the bars begin at the bottom at approximately minus \$4,200,000.”

# Specifically Ignoring Scale...



- Numbers have magnitude, not just order

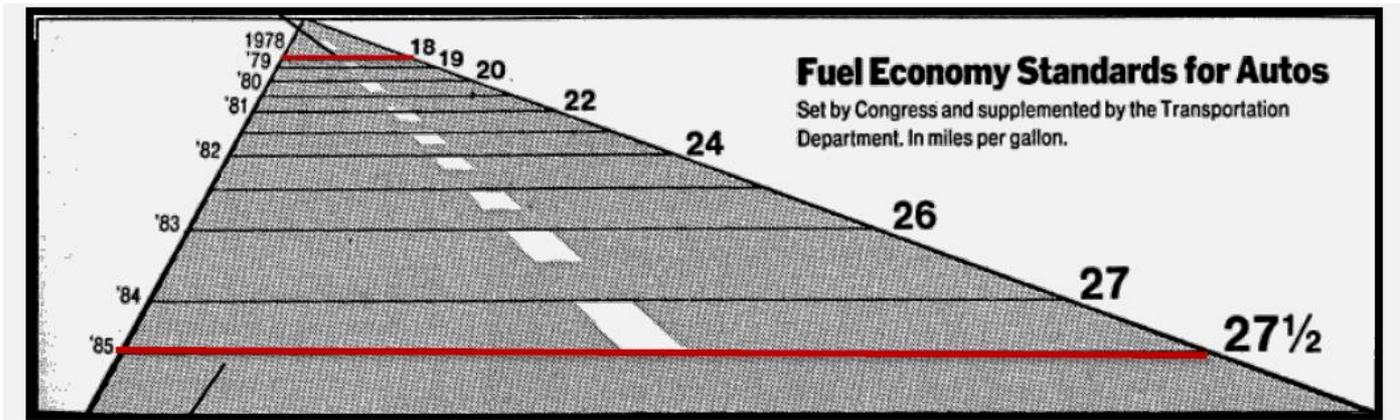
# The Lie Factor

- Tufte coined the term “lie factor” (LF), defined as

$$\text{Lie\_factor} = \frac{\text{size of graphic}}{\text{size of data}}$$

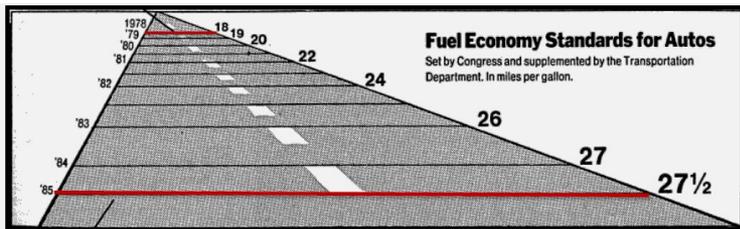
- High LF leads to exaggeration of differences or similarities
  - Deception
  - Misinterpretation
- For 0 to indicate not lying, what function could be applied to LF?

# Lie Factor Example



- Line representing 18 MPG in 1978 is 0.6 inches long
- Line representing 27.5 MPG in 1985 is 5.3 inches long

# Lie Factor Example



Ideally Lie Factor should be what?

- Increase in MPG from 1978 to 1985 is

$$\frac{27.5 - 18.0}{18.0} \times 100 = 53\%$$

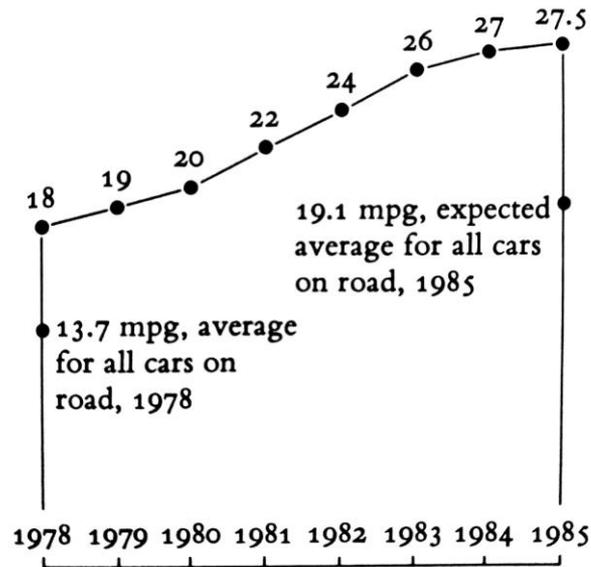
- Difference in length of lines is

$$\frac{5.3 - 0.6}{0.6} \times 100 = 783\%$$

- The Lie Factor is  $\frac{783}{53} = 14.8$

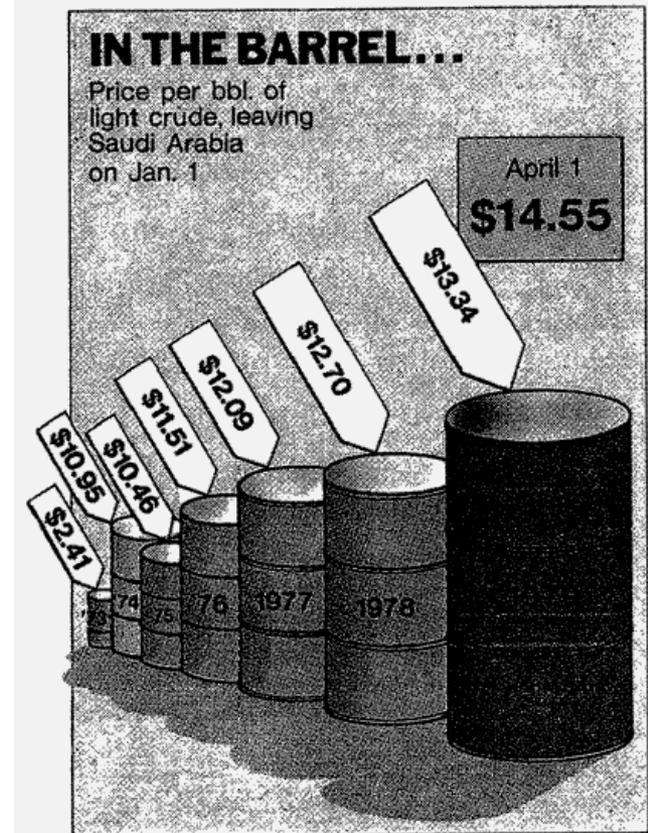
# Showing the Data without Lying

REQUIRED FUEL ECONOMY STANDARDS:  
NEW CARS BUILT FROM 1978 TO 1985



# Lie Factor Example

- LF of 9.4
- The LF is made worse by using a glyph with changing area as a visual representation of a linear change
- Price is keyed to glyph height
- But barrel area appears to change
  - ▣ Not just height



# Another Way to Lie: Design Variation

- The “3D” effect can be used to distort data
- 5 vertical scales:
  - 1973-1978: 1 inch = \$8
  - Jan-Mar 1 inch = \$4.73
  - Apr-Jun: 1 inch = \$4.37
  - Jul-Sep: 1 inch = \$4.16
  - Oct-Dec: 1 inch = \$3.92
- Two horizontal scales
  - 1973-1978: 1 inch = 3.8 years
  - 1979: 1 inch = 0.57 years



# Another Example of Area Mis-Used

- Both dimensions of the dollars are reduced by the same factor
- This squares the linear difference which the chart should show
- 1978 dollar should be twice as big....

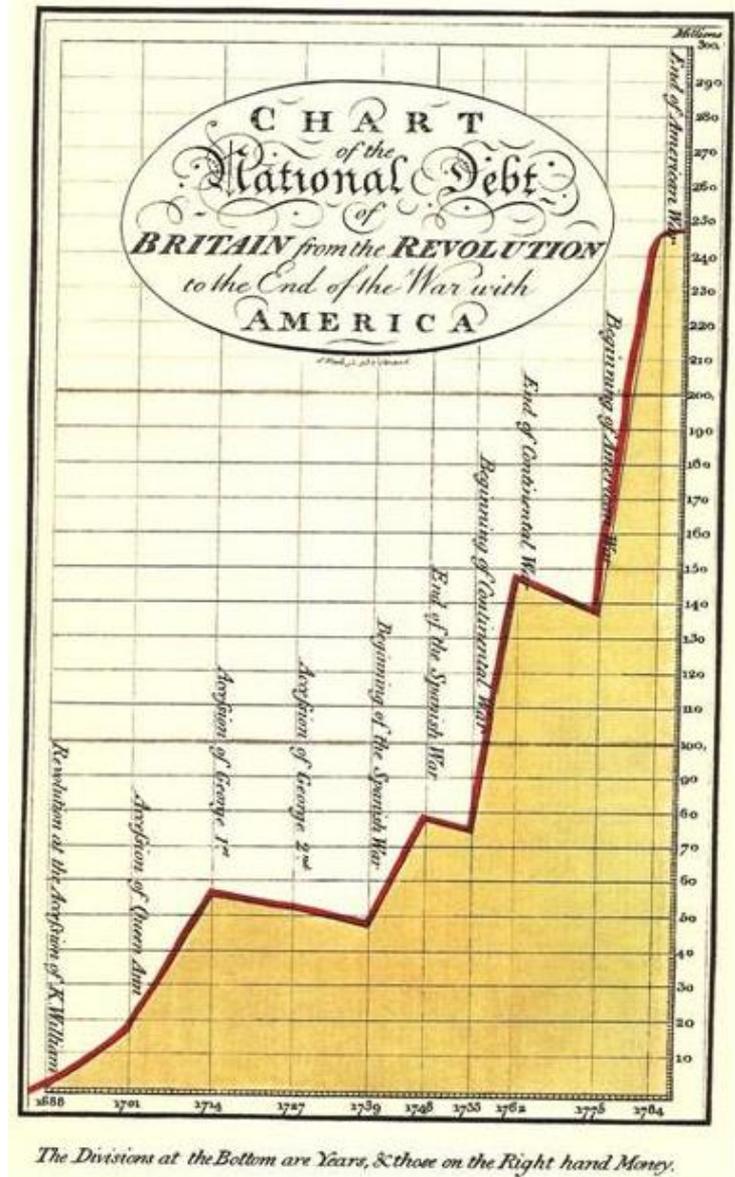


# How Does this Chart Distort the Data?

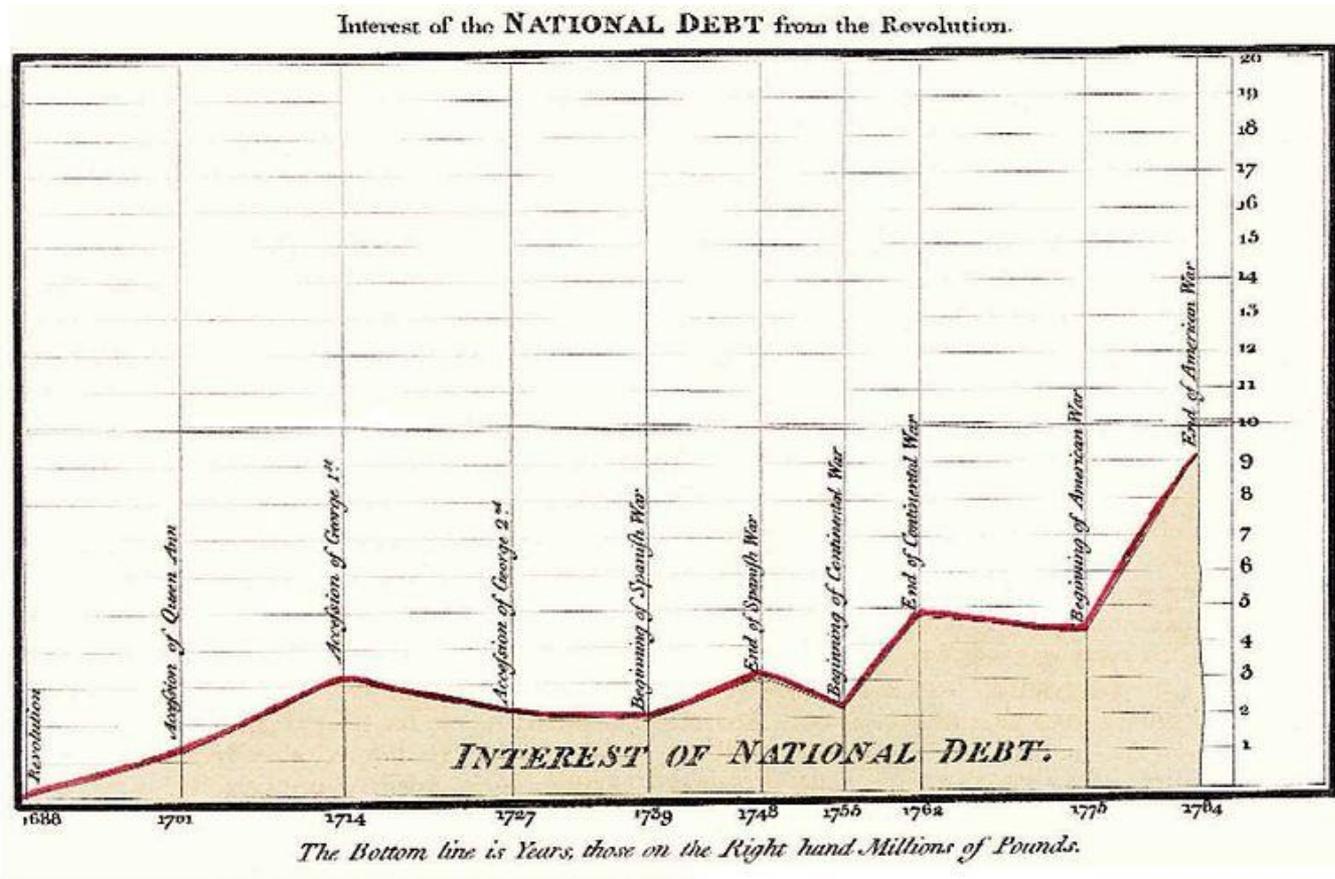
William Playfair 1786

Accompanied by a polemic decrying the financing of colonial wars with debt.

There are two issues...



# ...Just a Few Pages Later



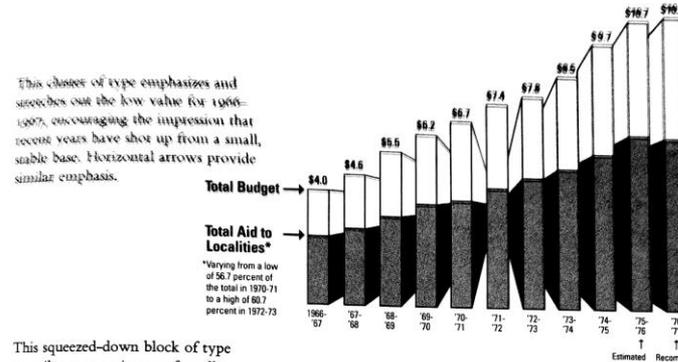
# Using Visual Elements to Deceive

The graphic wishes to display rising government spending

The budget, in fact, did not increase in the last 9 years shown

Some cues the chart provides:

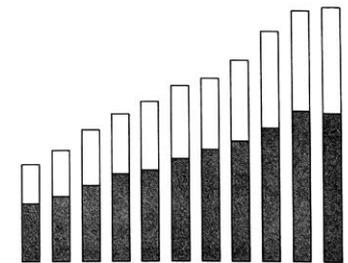
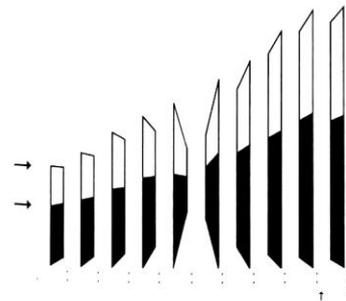
- Horizontal arrows on low values
- Vertical arrows on high values
- Angled 3D-effect lines to provide impression of increase



This squeezed-down block of type contributes to an image of small, squeezed-down budgets back in the good old days.

Arrows pointing straight up emphasize recent growth. Compare with horizontal arrows at left.

Leaving behind the distortion in the chartjunk heap at the left yields a calmer view:



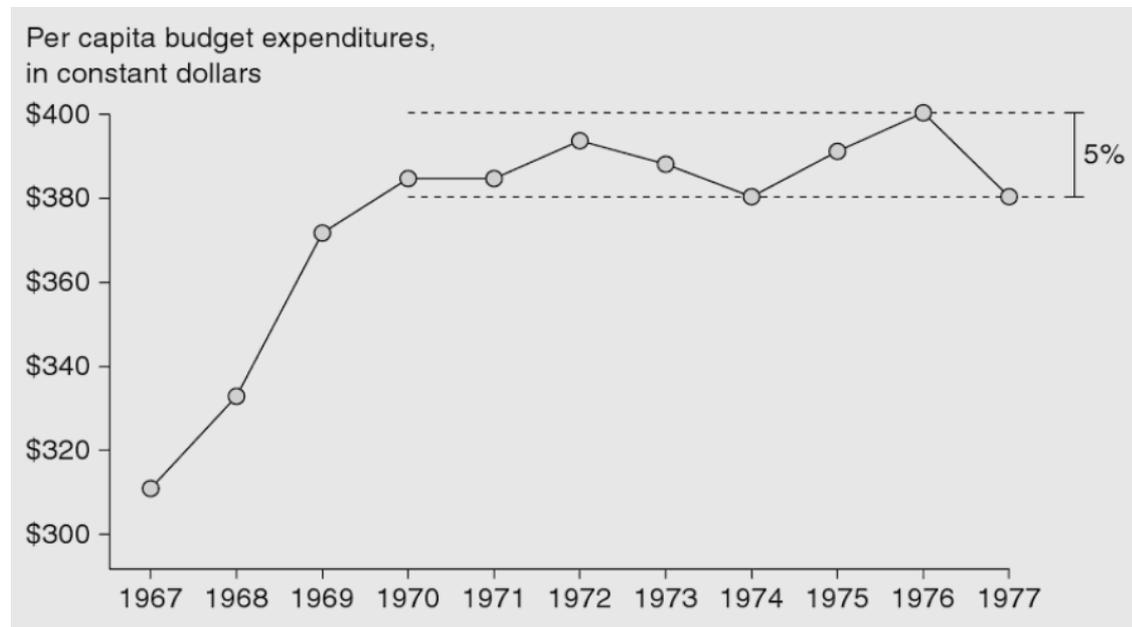
# Removing Statistical Lapses as well

In addition to visual cues:

State population increase neglected in original graphic

Original graphic does not adjust for inflation

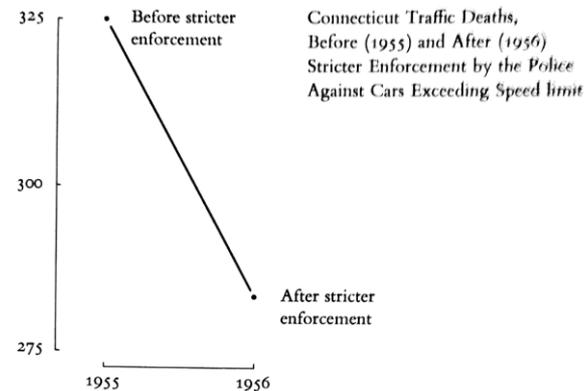
Figure at right shows budget accounting for those issues



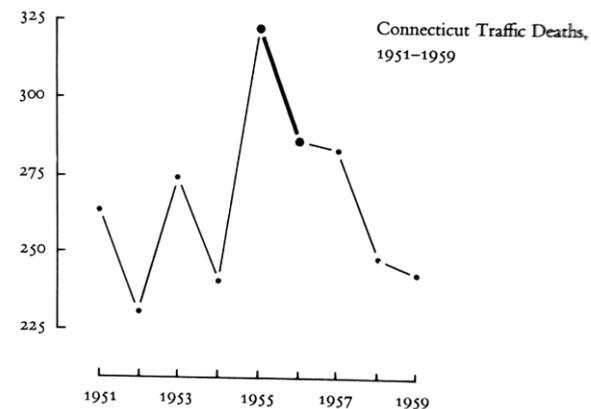
# Lying by Removing Context

Removing extra data points leads one to obvious conclusion that stricter speed limit enforcement reduced deaths.

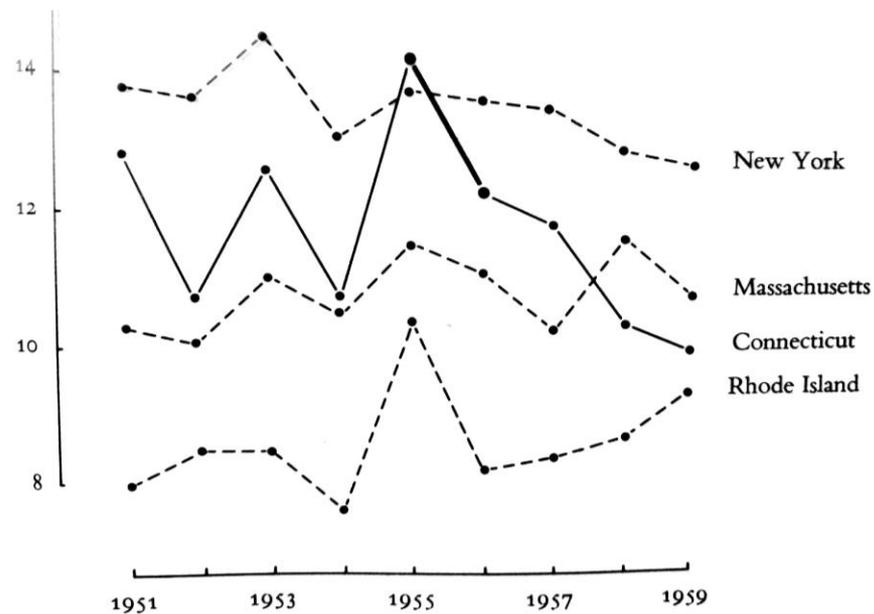
How does the addition of more sample points make that conclusion questionable?



A few more data points add immensely to the account:



# ...Adding More Context



How does the addition of data from other states effect the interpretation of the cause for the decline in Connecticut?

# One Last Example...



PRAVDA, MAY 24, 1982, p.2.

What's deceptive here?

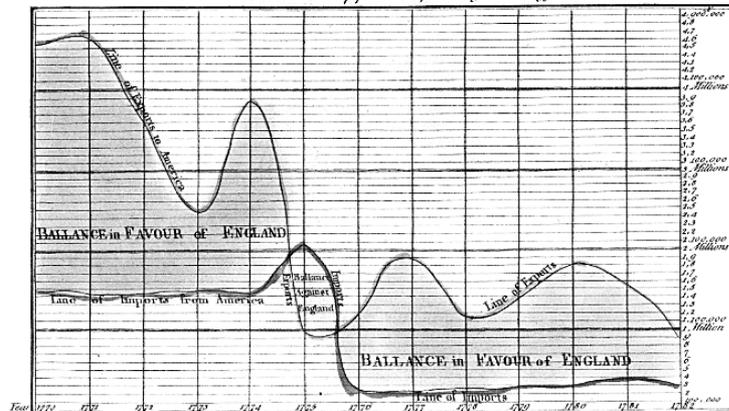
# Good Design Practice

Tufte suggests a ratio of data ink to total ink be close to 1

Data ink is ink devoted to non-redundant display of data

Observe Playfair's evolution:

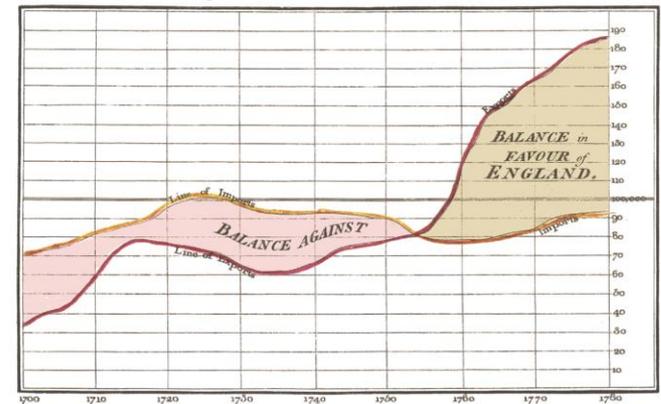
CHART of IMPORTS and EXPORTS of ENGLAND to and from all NORTHAMERICA  
From the Year 1770 to 1782 by W. Playfair



The Bottom Line is divided into Years the right hand Line into HUNDRED THOUSAND POUNDS  
Published in the Advertiser 20<sup>th</sup> Aug<sup>r</sup> 1786.

J. Smith Sculp<sup>r</sup>

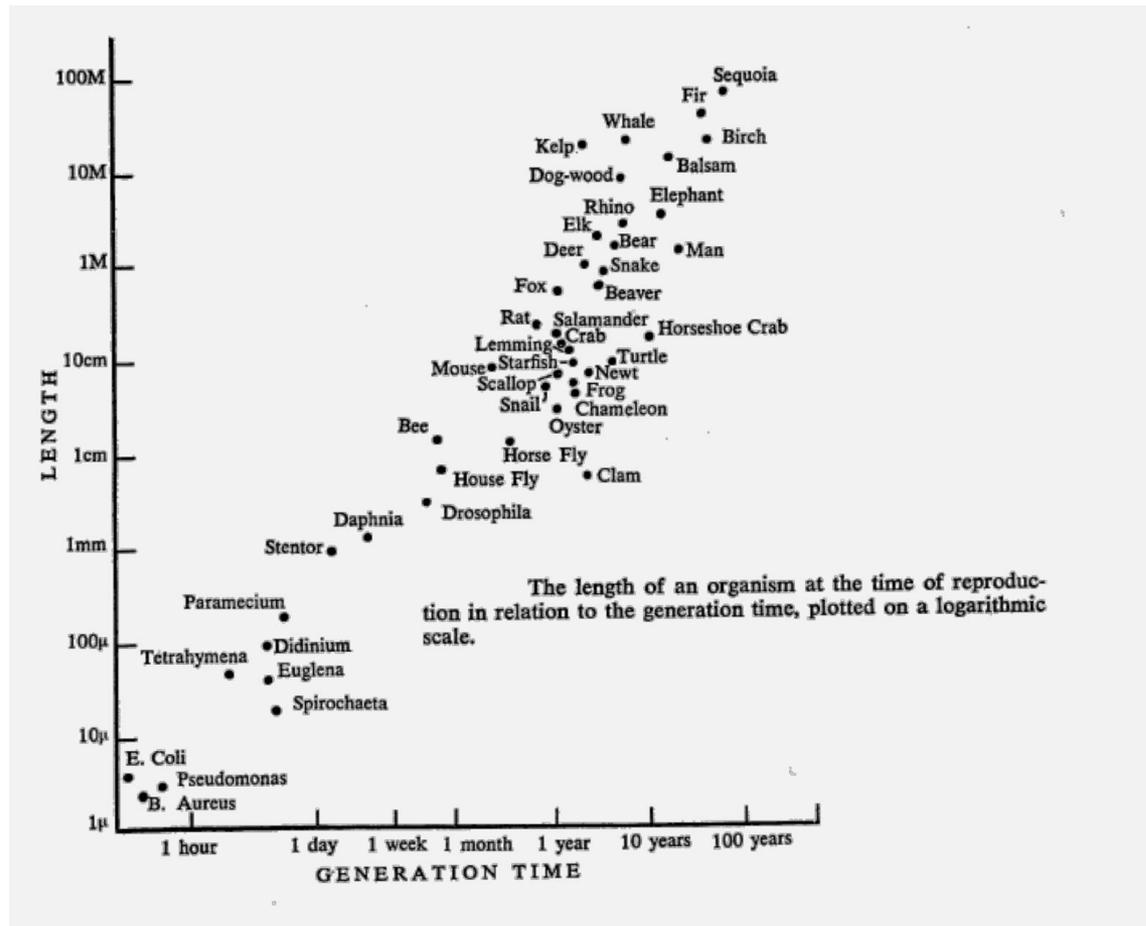
Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780



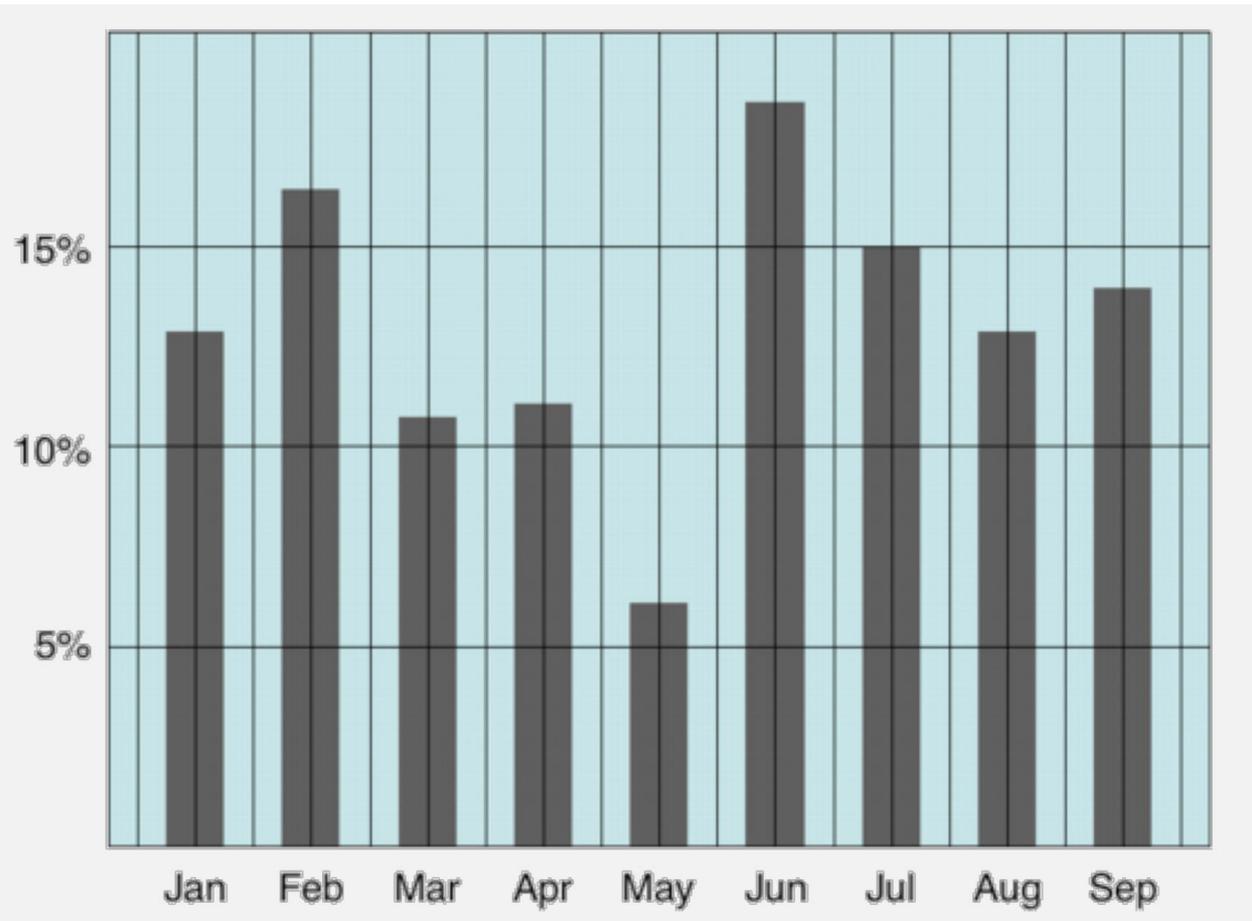
The Bottom line is divided into Years, the Right hand line into L10,000 each.  
Published in the Advertiser, 14<sup>th</sup> May 1786, by W. Playfair.

Printed and Sold by J. Smith, Strand, London.

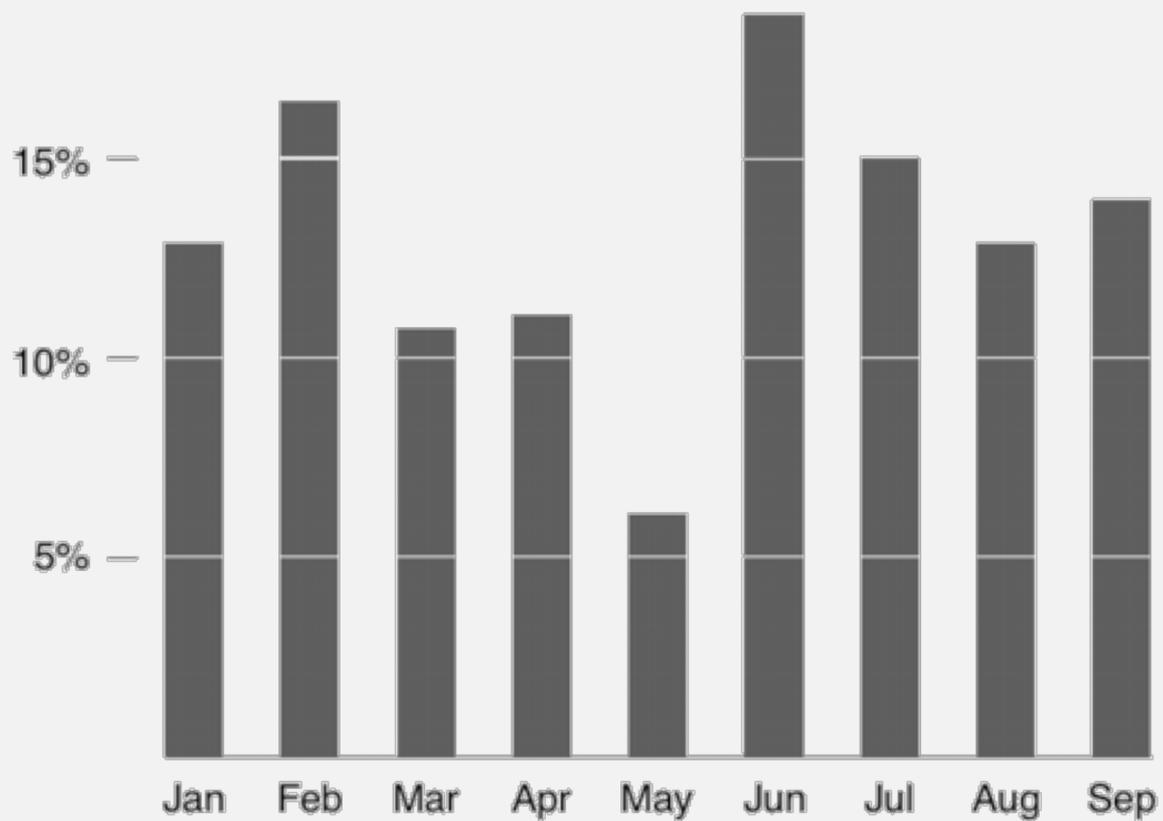
# Good....



# Less Good...

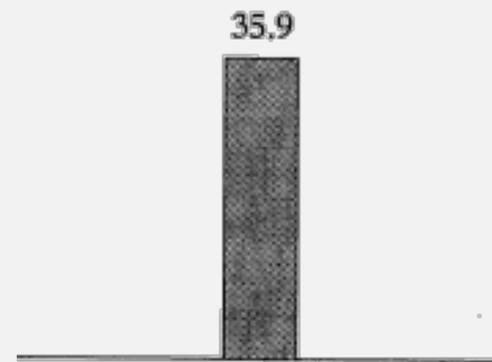


# Improved



# Erasing Non-Data Ink

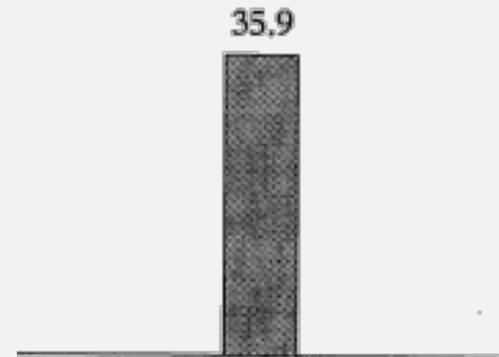
How many times is height encoded?



# Erasing Non-Data Ink

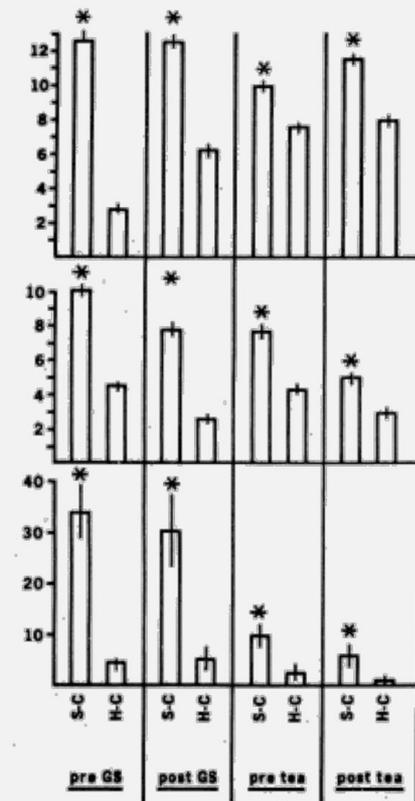
## Multiple encodings:

1. Height of the left line
2. Height of the right line
3. Height of shading
4. Position of top horizontal line
5. Position (placement) of the number
6. Value of the number



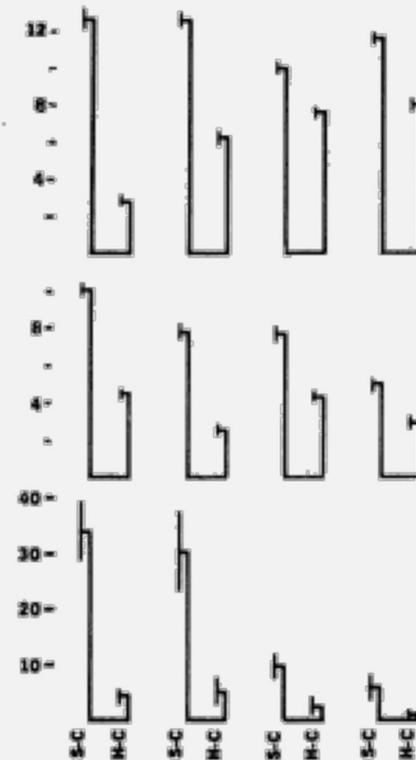
# Erasing Non-Data Ink: Example

Results of a study indicating that one type of element always has a higher value under different experimental conditions

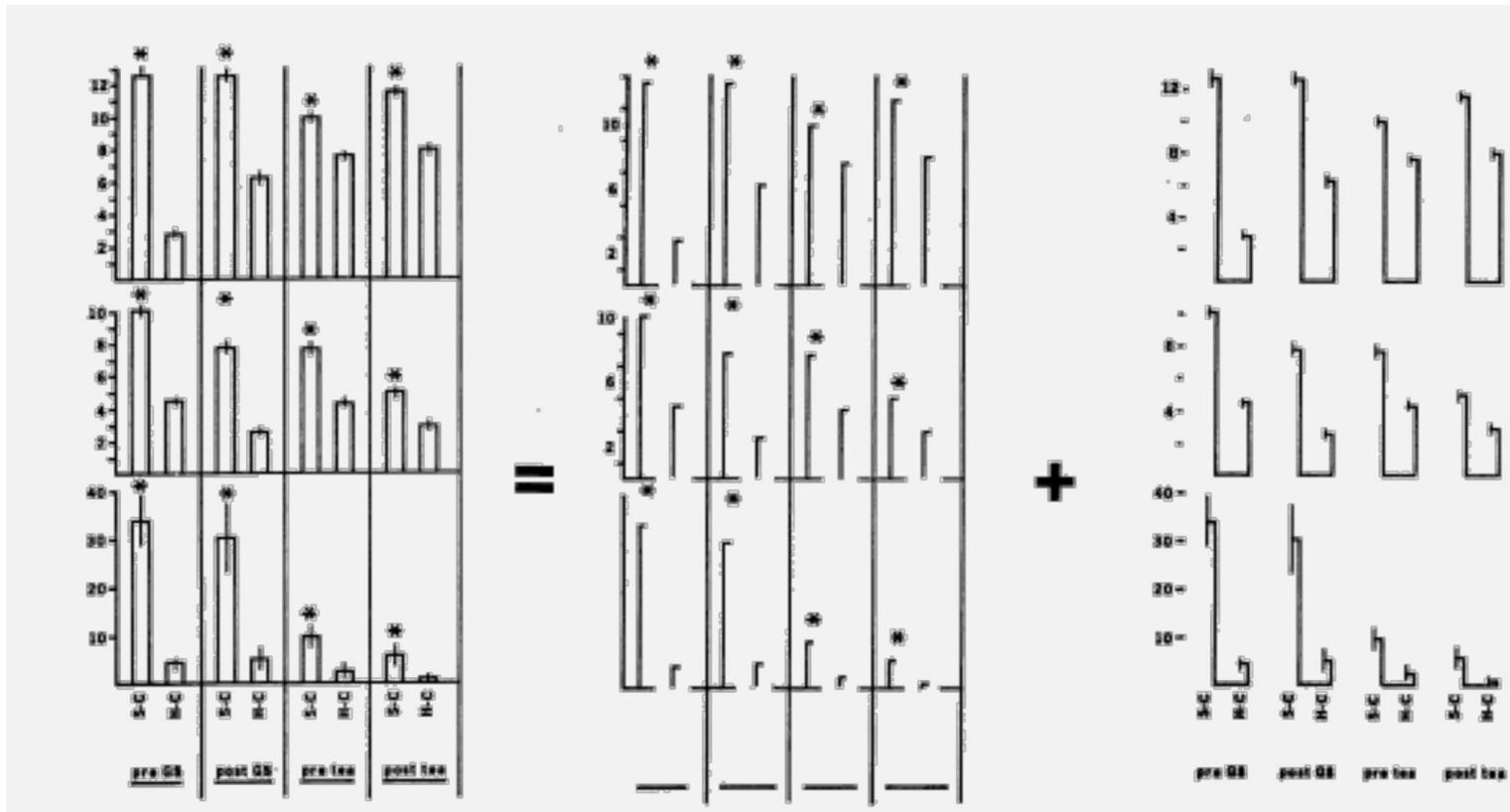


# After Removing Non-Data Ink

After removing all non-data ink



# Is the Data Easier to Interpret?



# Similarly....



# Tufte: Principles of Graphical Excellence

Graphical excellence is the well-designed presentation of interesting data – a matter of substance, of statistics, and of design

Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency.

Graphical excellence is that which gives the viewer the greatest number of ideas in the shortest time with the least ink in the smallest place.

Graphical excellence is nearly always multivariate.

Graphical excellence requires telling the truth about the data.

# Is Tufte Correct?

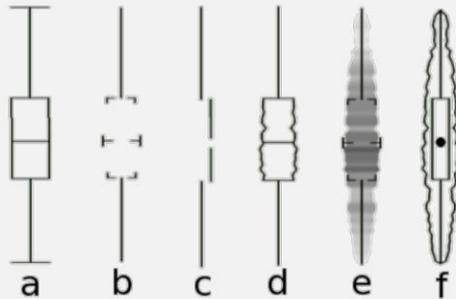
## A User Study of Visualization Effectiveness Using EEG and Cognitive Load

E. W. Anderson<sup>1</sup>, K. C. Potter<sup>1</sup>, L. E. Matzen<sup>2</sup>, J. F. Shepherd<sup>2</sup>, G. A. Preston<sup>3</sup>, and C. T. Silva<sup>1</sup>

<sup>1</sup>SCI Institute, University of Utah, USA

<sup>2</sup>Sandia National Laboratories, USA

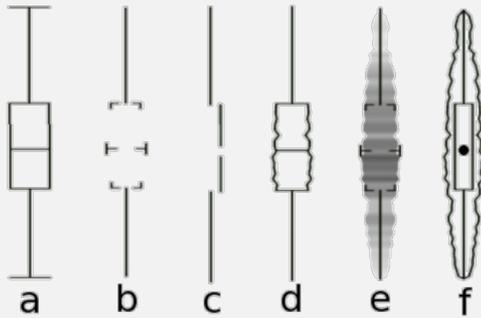
<sup>3</sup>Utah State Hospital, USA



- Asked participants to choose the box plot with the largest range from a set
- Varied representations
- Measured cognitive load from EEG brain waves



# Results



The simplest box plot is the hardest to interpret

# Is ChartJunk Useful?

## Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

Scott Bateman, Regan L. Mandryk, Carl Gutwin,  
Aaron Genest, David McDine, Christopher Brooks

Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada  
scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca,  
aaron.genest@usask.ca, dam085@mail.usask.ca, cab938@mail.usask.ca

### ABSTRACT

Guidelines for designing information charts often state that the presentation should reduce ‘chart junk’ – visual embellishments that are not essential to understanding the data. In contrast, some popular chart designers wrap the presented data in detailed and elaborate imagery, raising the questions of whether this imagery is really as detrimental to understanding as has been proposed, and whether the visual embellishment may have other benefits. To investigate these issues, we conducted an experiment that compared embellished charts with plain ones, and measured both interpretation accuracy and long-term recall. We found that people’s accuracy in describing the embellished charts was no worse than for plain charts, and that their recall after a two-to-three-week gap was significantly better. Although we are cautious about recommending that all charts be produced in this style, our results question some of the premises of the minimalist approach to chart design.

### Author Keywords

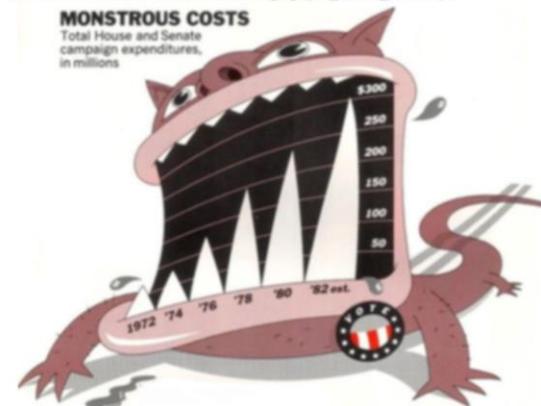
Charts, information visualization, imagery, memorability.

### ACM Classification Keywords

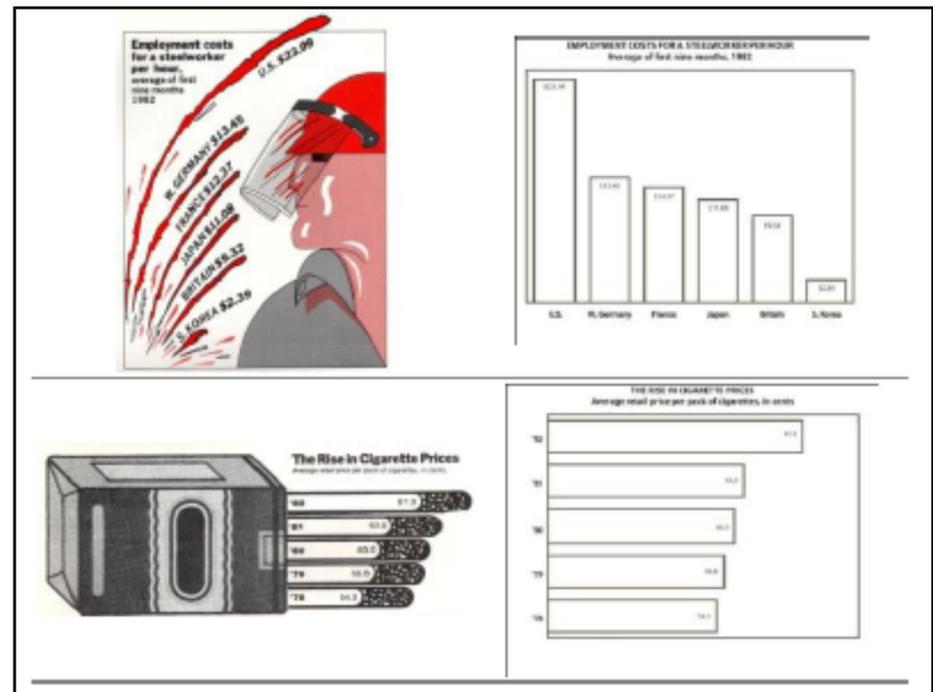
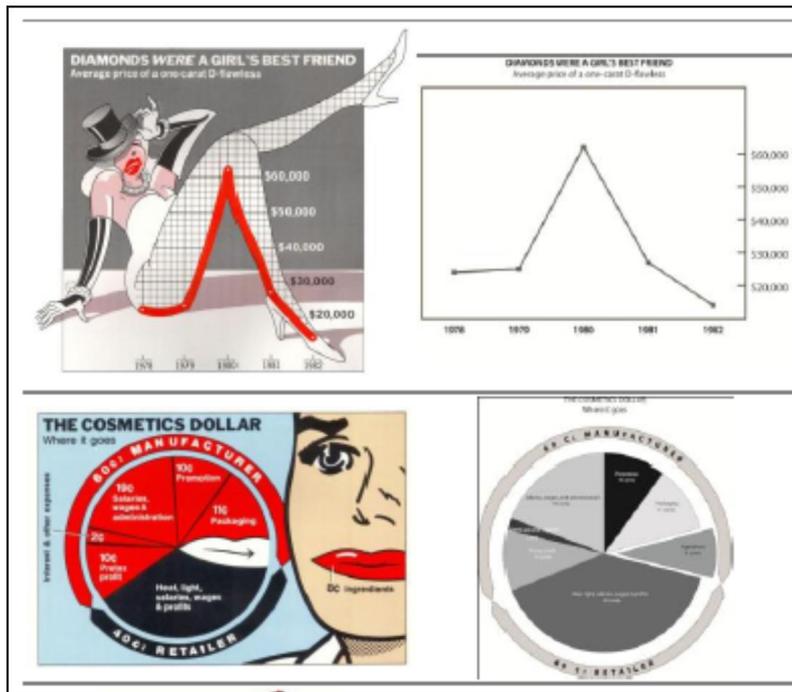
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

*data-ink* – or the ink in the chart used to represent data.

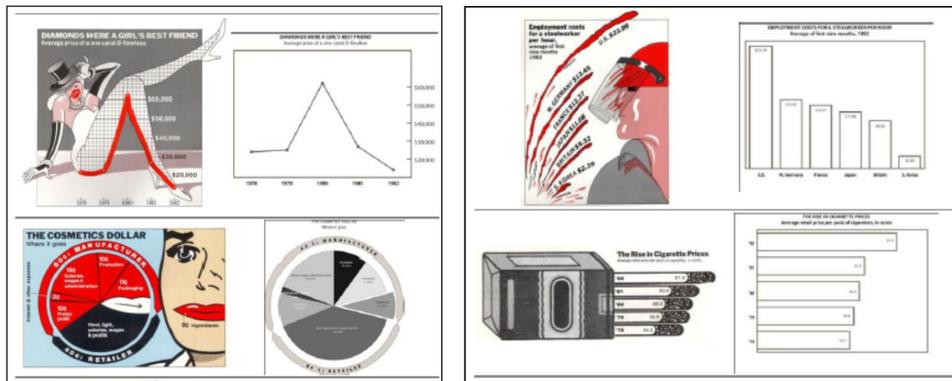
Despite these minimalist guidelines, many designers include a wide variety of visual embellishments in their charts, from small decorations to large images and visual backgrounds. One well-known proponent of visual embellishment in charts is the graphic artist Nigel Holmes, whose work regularly incorporates strong visual imagery into the fabric of the chart [7] (e.g., Figure 1).



# Redesigning Holmes Charts



# Results



1. **No significant difference** between **interpretation accuracy**
2. **No significant difference** in **recall accuracy** after a five-minute gap
3. **Significantly better recall** for Holmes charts of both chart topic and the details (categories and trend) after long-term gap (2-3 weeks).
4. Participants found the Holmes charts **more attractive, more enjoyable**, and were **easiest and fastest to remember**.