

Design Principles: Tufte, Few, and others

CS 4460 - Information Visualization
Spring 2019
Alex Endert



Alex Endert

endert@gatech.edu

Today's Agenda

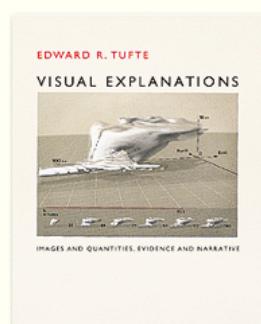
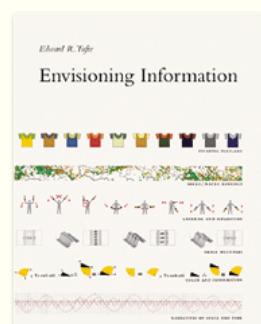
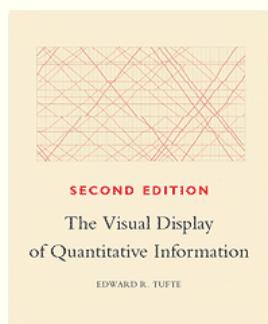
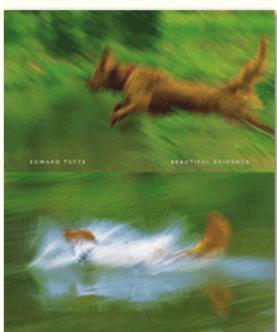


Edward Tufte has written seven books, including *Visual Explanations*, *Envisioning Information*, *The Visual Display of Quantitative Information*, and *Data Analysis for Politics and Policy*. He writes, designs, and self-publishes his books on analytical design, which have received more than 40 awards for content and design. He is Professor Emeritus at Yale University, where he taught courses in statistical evidence, information design, and interface design. His current work includes landscape sculpture, printmaking, video and a new book.

This website describes Edward Tufte's books, one-day course, and artwork. For further information, call Graphics Press at 203 272-9187, or fax 203 272-8600, or [email](#).

For a moderated forum on analytical design, go to [ASK E.T.](#)

BOOKS



Graphical Excellence

- Principles

- 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**
- **Telling the truth** about the data

According to Tufte

Leveraging Human Capabilities

- Data graphics should complement what humans do well

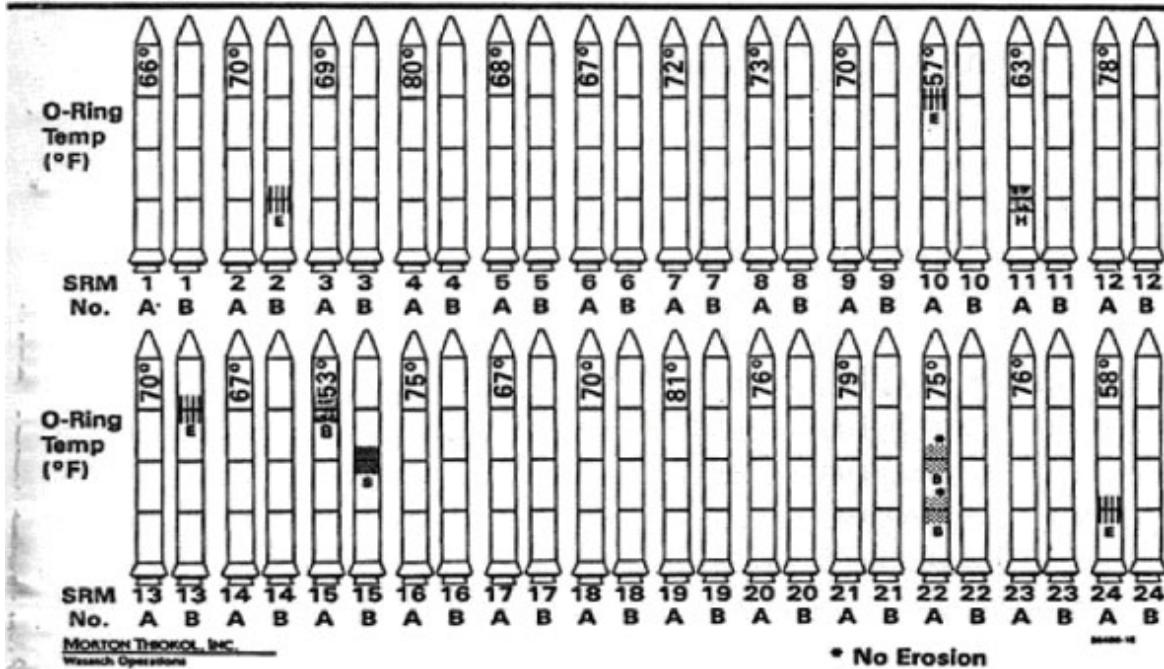
“We thrive in information-thick worlds because of our marvelous and everyday capacities to *select, edit, single out, focus, organize, condense, reduce, boil down, choose, categorize, catalog, classify, list, abstract, scan, look over, sort, integrate, blend, inspect, filter, lump, skip, smooth, chunk, average, approximate, cluster, aggregate, outline, summarize, itemize, review, dip into, flop through, browse, glance into, leaf through, skim, refine, enumerate, glean, synopsize, winnow the wheat from the chaff, and separate the sheep from the goats.*” Vol.2, page 50

:)

Summary

- 1. Tell the truth baseline? scale? context? (Any limitation of what the data can tell us about the situation?)
 - Graphical integrity
- 2. Do it effectively with clarity, precision...
 - Design aesthetics

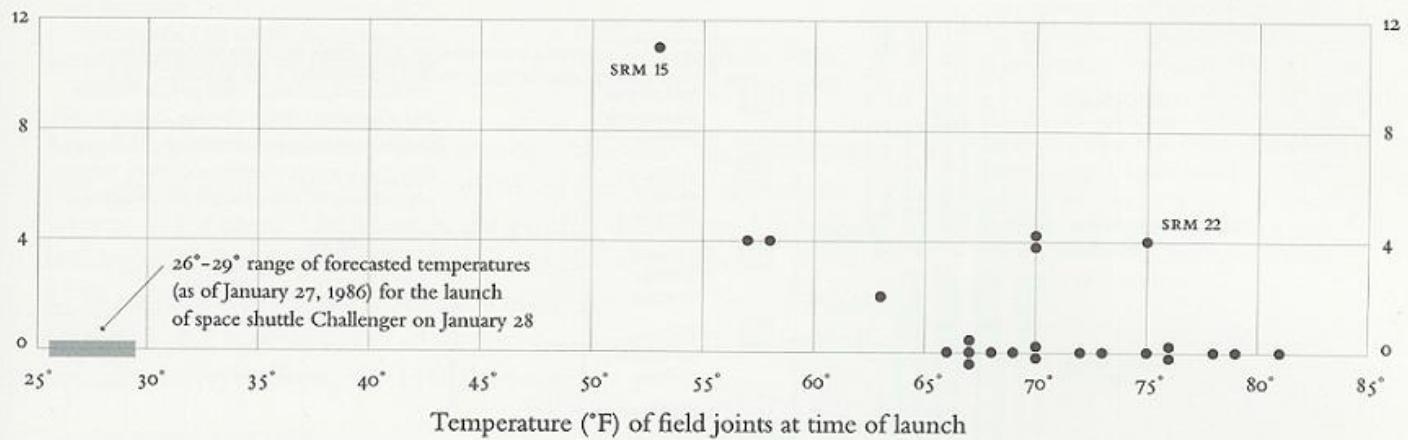
History of O-Ring Damage in Field Joints (Cont)



* No Erosion

INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION
AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

O-ring damage
index, each launch



Again...

- 1. Tell the truth
 - Graphical integrity
- 2. Do it effectively with clarity, precision...
 - Design aesthetics

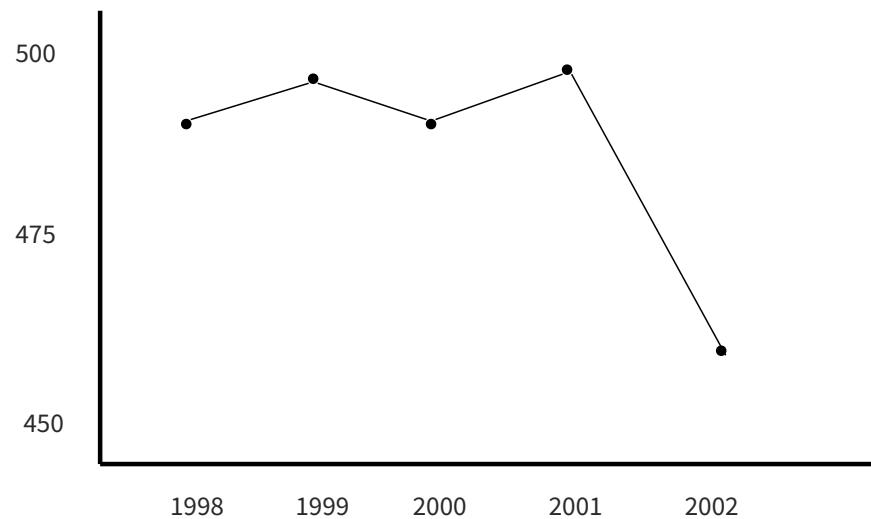
Let's look at each of these in more detail

1. Graphical Integrity

- Your graphic should tell the truth about your data

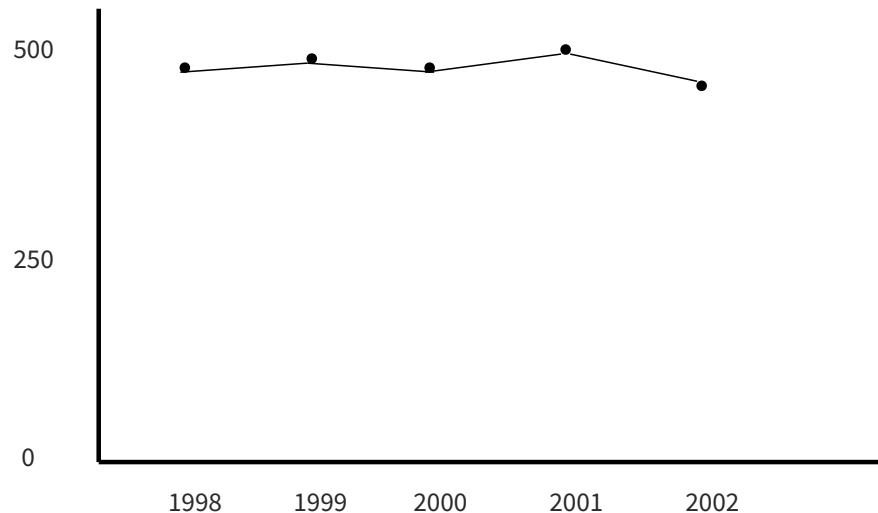
Example

Stock market crash?



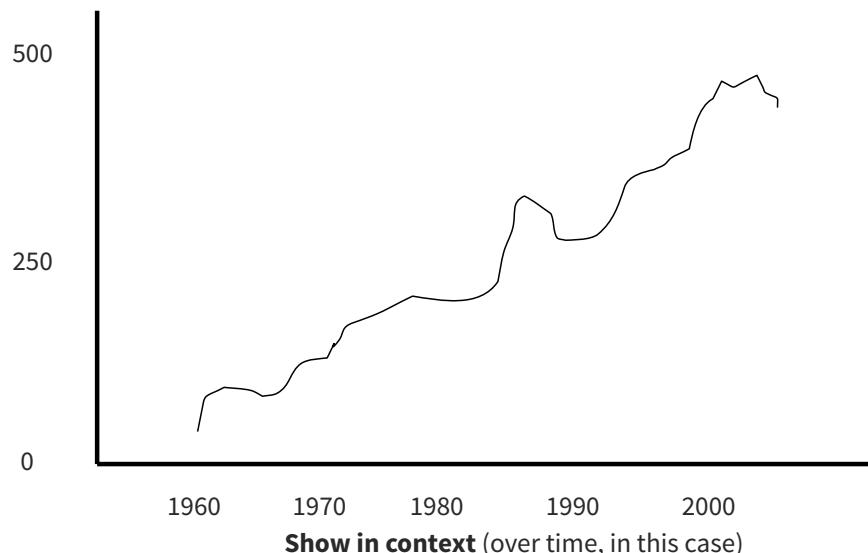
Example

Show entire scale



Example

* for static visuals, interactivity can allow for zooming into an area



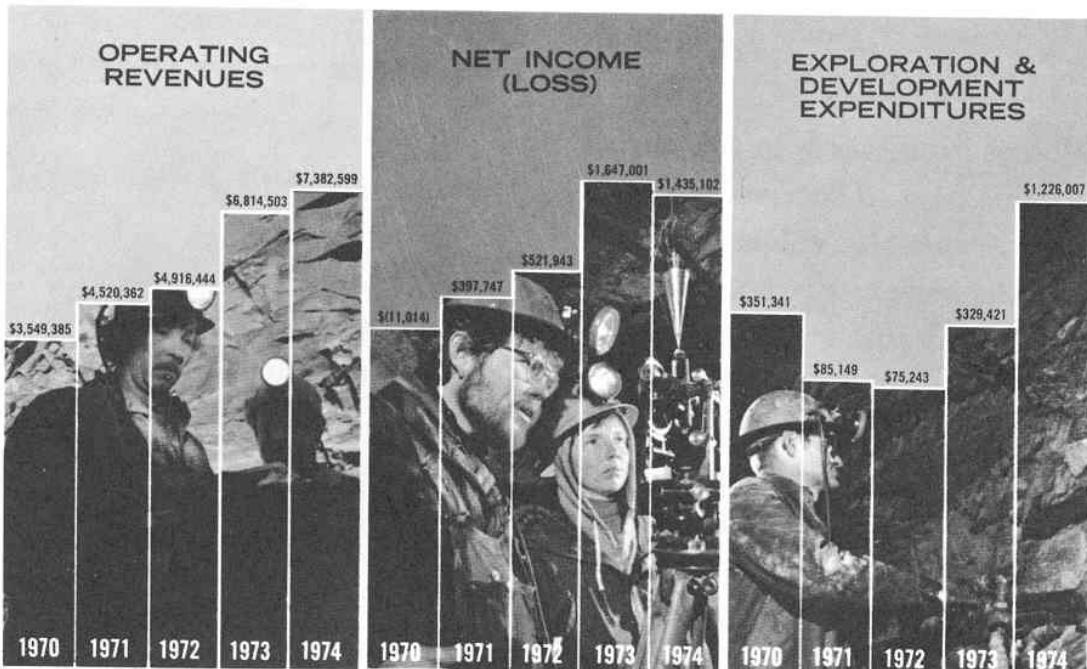
Show in context (over time, in this case)

Chart Integrity Principles

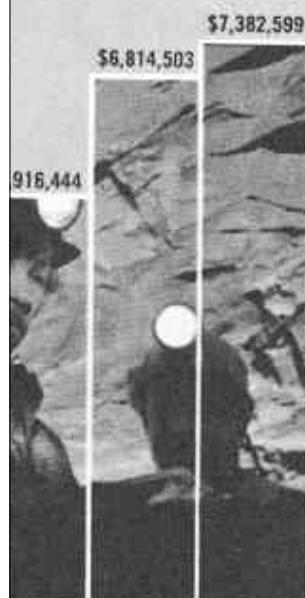
- Where's the **baseline**?
- What's the **scale**?
- What's the **context**?

Vol 1, p. 54 (1)

What are the problems here?



RATING VENUES

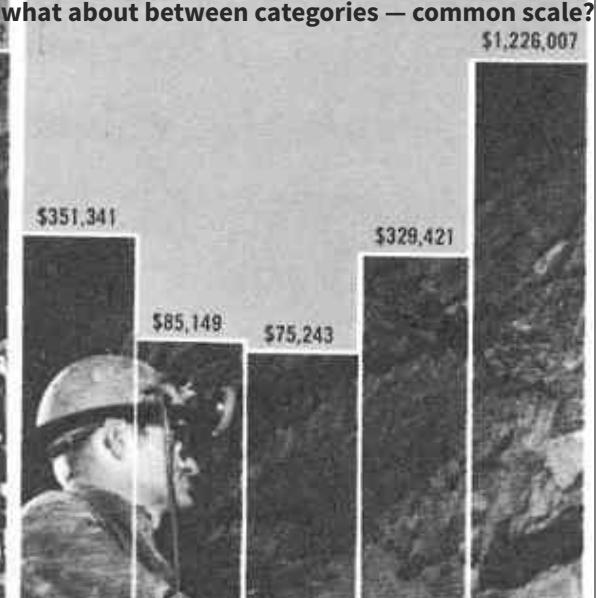


NET INCOME (LOSS)



EXPLORATION & DEVELOPMENT EXPENDITURES

Where's O?
Note middle '70
what about between categories — common scale?



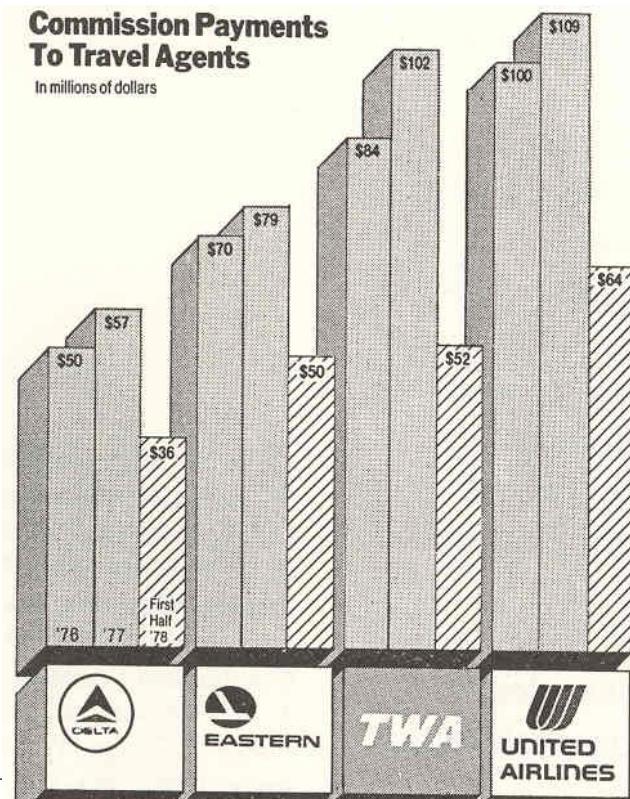
Vol 1, p 54 (2)

What's being compared?

Note how hard it is to judge if the bars with height \$50 are equal

Commission Payments To Travel Agents

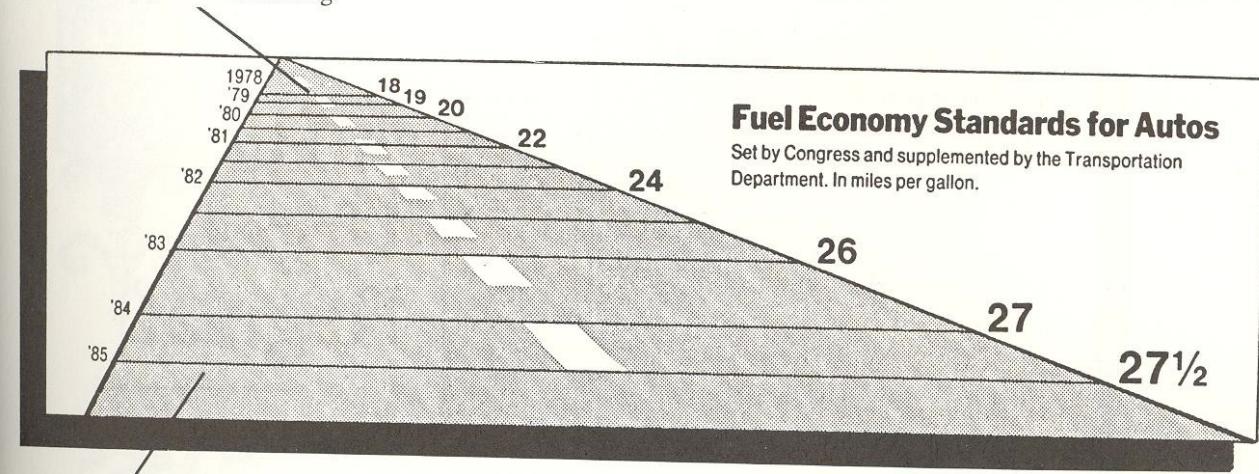
In millions of dollars



Vol 1, 57

Scale?

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



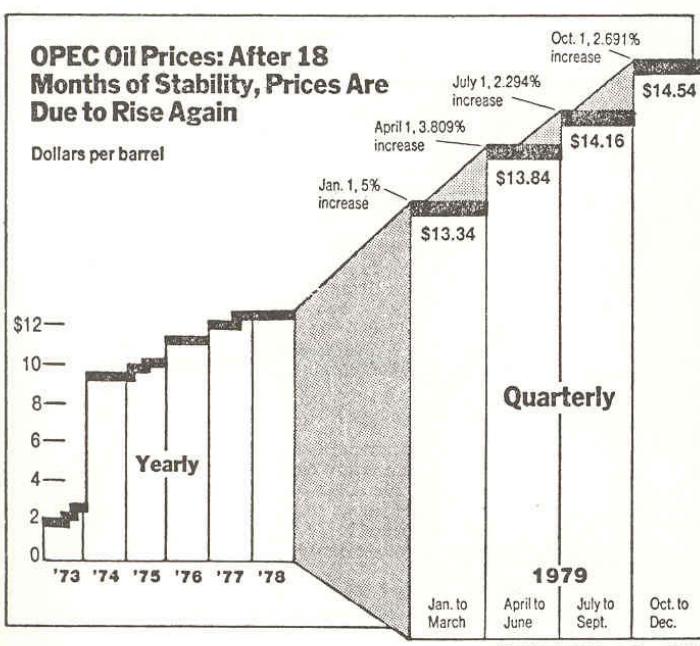
This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

Vol 1, p. 61

Scale?

OPEC Oil Prices: After 18 Months of Stability, Prices Are Due to Rise Again

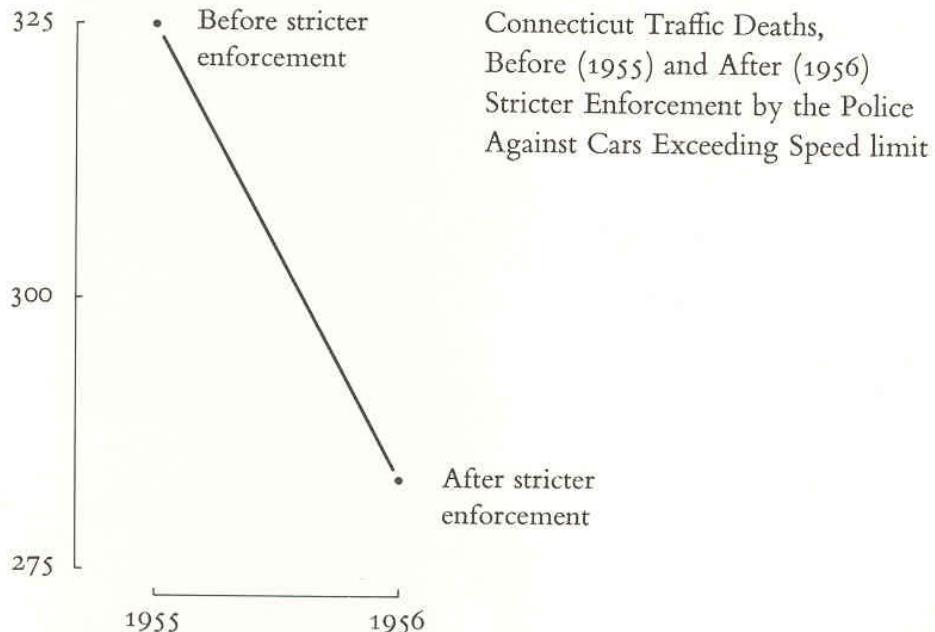
Dollars per barrel



New York Times, December 19, 1978,
p. D-7.

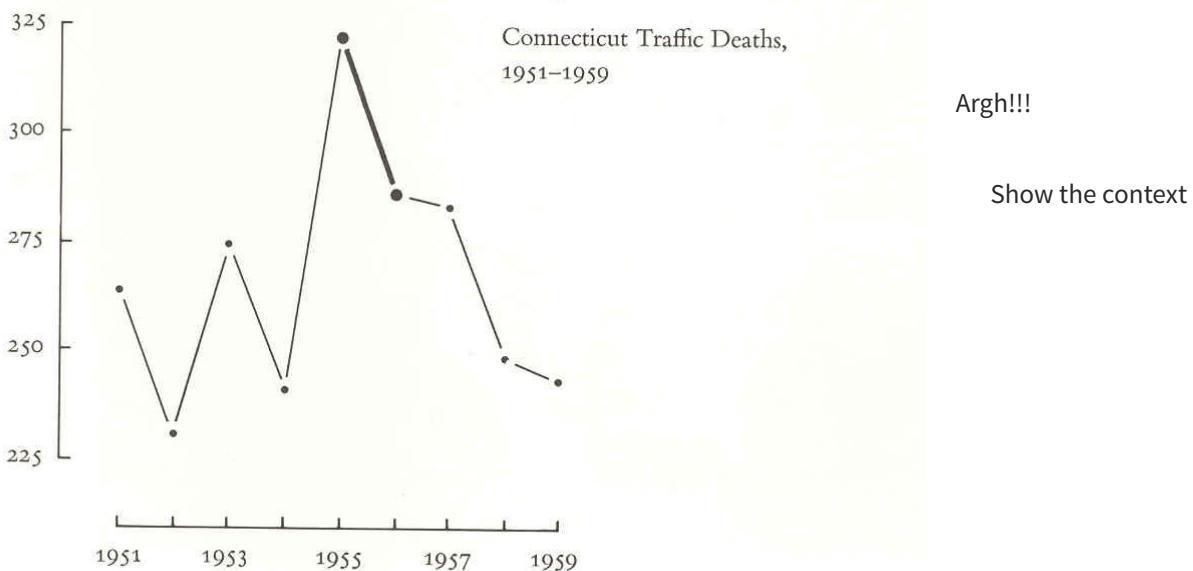
Vol 1, p. 74

Great work! ... right?



Vol 1, p. 74

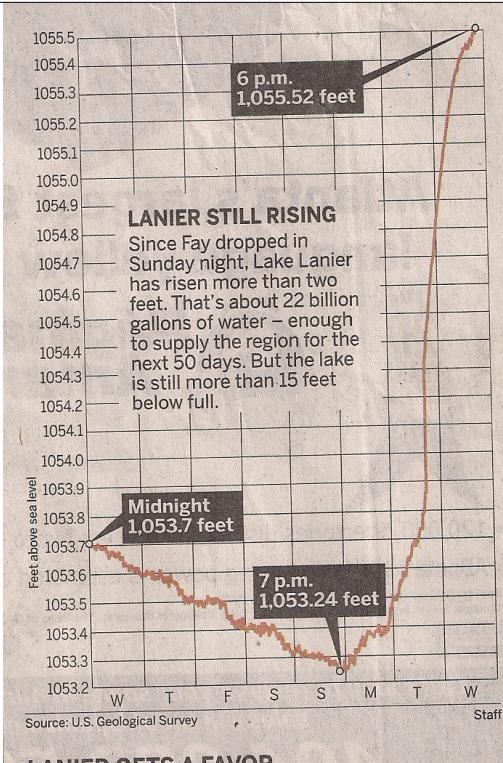
A few more data points add immensely to the account:



Local Example

A huge rise?

Atlanta Journal Constitution
Summer '08



endert@gatech.edu

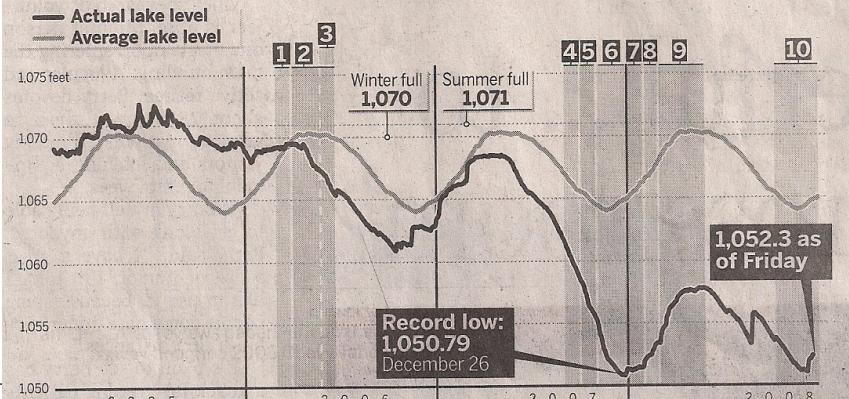
More of the data

Atlanta Journal
Constitution
Dec. '08

LANIER STILL ON LIFE SUPPORT

Recent rain helps Lake Lanier, but metro Atlanta's primary water source is a long way from normal.

- 1 March 2006: Drought begins.
- 2 April - June 2006: Faulty gauge leads to two-foot drop.
- 3 June 2006: Drought officially declared, triggering statewide watering restrictions. State officials warn the metro region's water supply is at risk.
- 4 September 2007: State enacts near-total ban on outdoor watering in North Georgia. Federal officials warn that Lanier is likely to hit a new record low.
- 5 October 2007: State predicts Lanier could "run dry" in 80 days and mandates 10 percent cut in North Georgia's water use.
- 6 November - December 2007: The White House brokers a deal with Georgia, Alabama and Florida to keep more water in Lanier.
- 7 January 2008: General Assembly passes statewide water plan, but with no new funds.
- 8 February 2008: Gov. Perdue eases restrictions on outdoor water use to allow some landscape watering and filling of swimming pools.
- 9 March - May 2008: The U.S. Army Corps of Engineers cuts water released down the Chattahoochee River by 13 percent, holding more in Lanier.
- 10 November 2008 and continuing through April 2009: The corps minimizes water released from Lanier.



endert@gatech.edu

Watch Size Coding

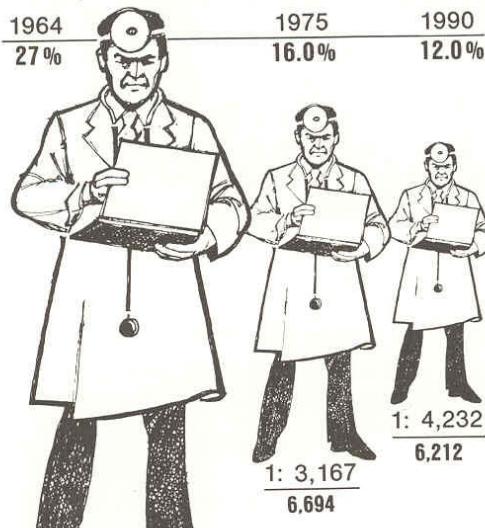
- Height/width vs. area vs. volume

Vol 1, p. 69

THE SHRINKING FAMILY DOCTOR In California

Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27 %	16.0 %	12.0 %



area = value?

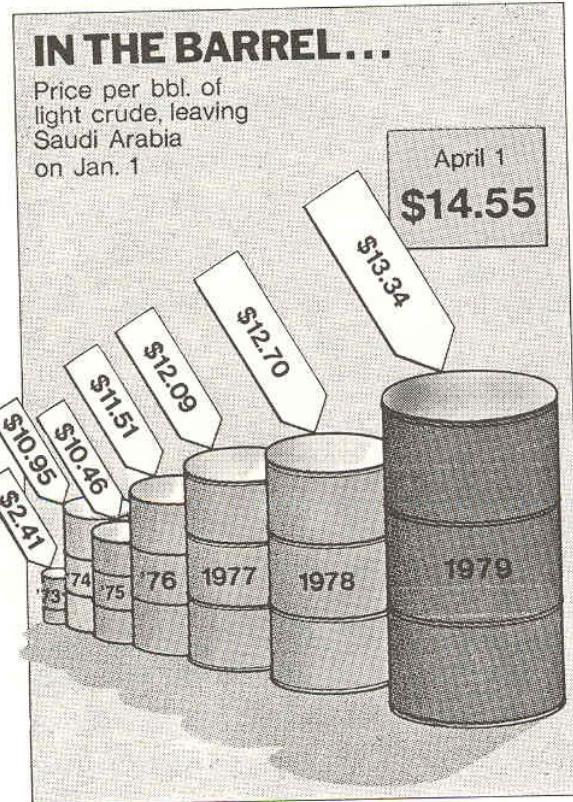
height? width?

1: 2,247 RATIO TO POPULATION
8,023 Doctors

Los Angeles Times

Vol 1, p. 62

volume = value?



Georgia Tech

Measuring Misrepresentation

- Visual attribute value should be directly proportional to data attribute value
- Potential way to measure this:
 - How far off 1:1 ratio?

$$\text{Lie factor} = \frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}}$$

Georgia Tech

Alex Endert

endert@gatech.edu

2. Design Aesthetics

- Set of principles to help guide designers

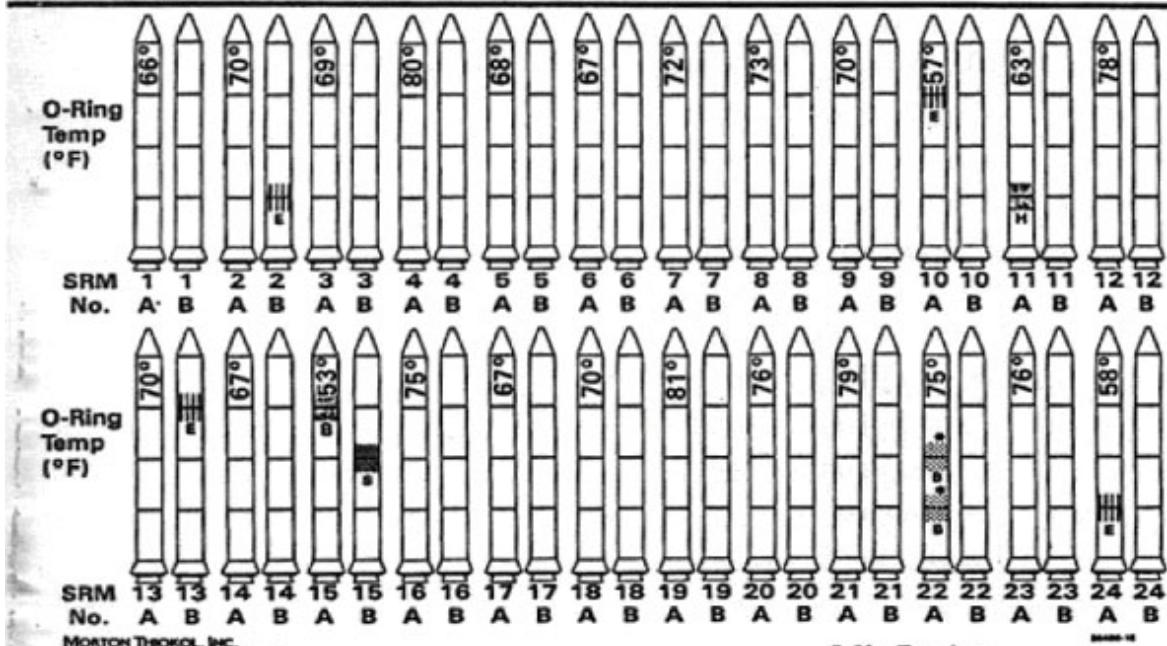
Design Principle

- Maximize **data-ink ratio**

$$\text{Data ink ratio} = \frac{\text{Data ink}}{\text{Total ink used in graphic}}$$

= proportion of graphic's ink devoted
to the non-redundant display of
data-information

History of O-Ring Damage in Field Joints (Cont)



* No Erosion
Focus ink on display of non-redundant part.
Avoid too much superfluous information

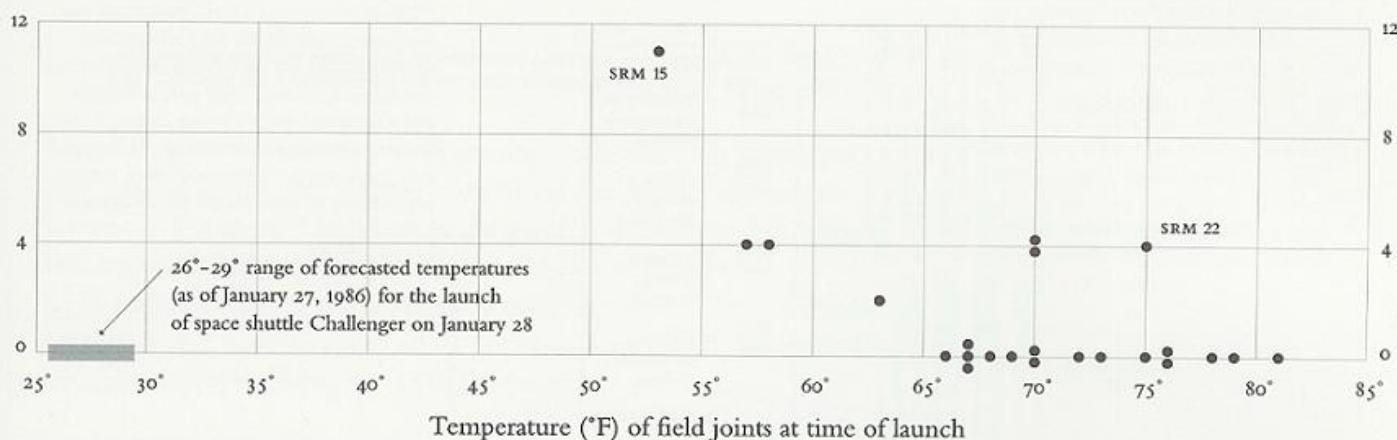
MORTON THOKOL, INC.
Research Operations

INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION
AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

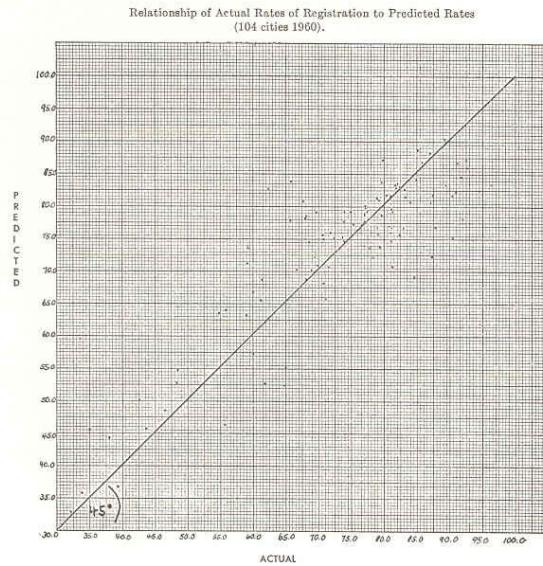
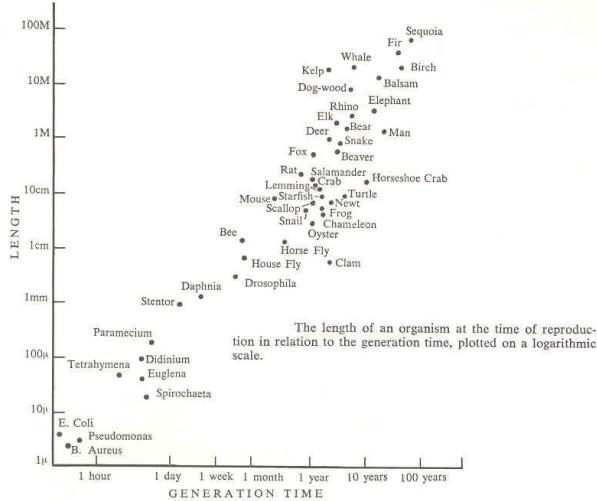
Alex Endert

endert@gatech.edu

O-ring damage
index, each launch



Vol 1, p. 94



Good

Bad

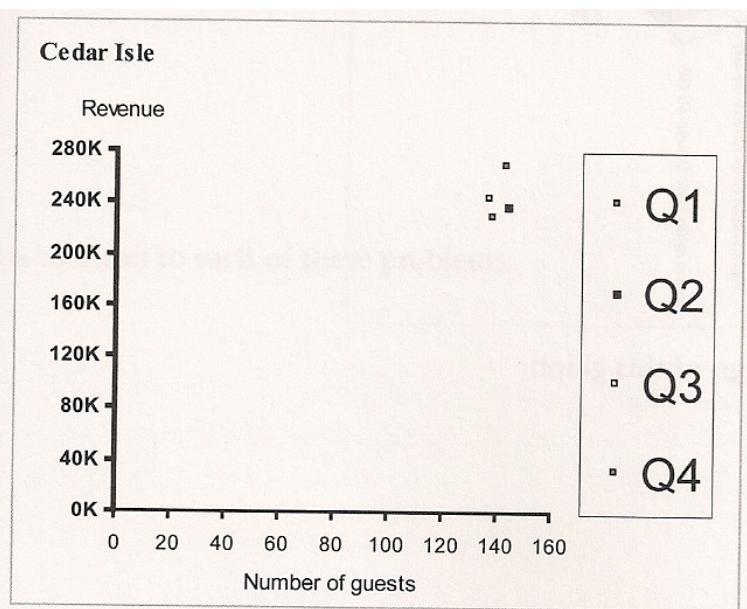
More...

- Above all else, **show the data**
- Maximize the **data-ink ratio**
- Erase non-data-ink
- Erase redundant data-ink
- **Revise** and edit

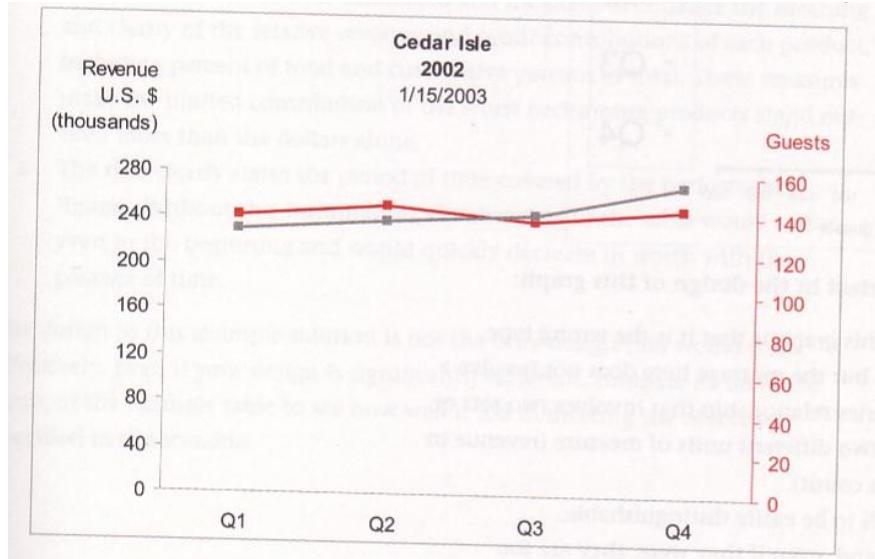
Examples of redesigning charts

- Bar chart, scatter plot, box plot

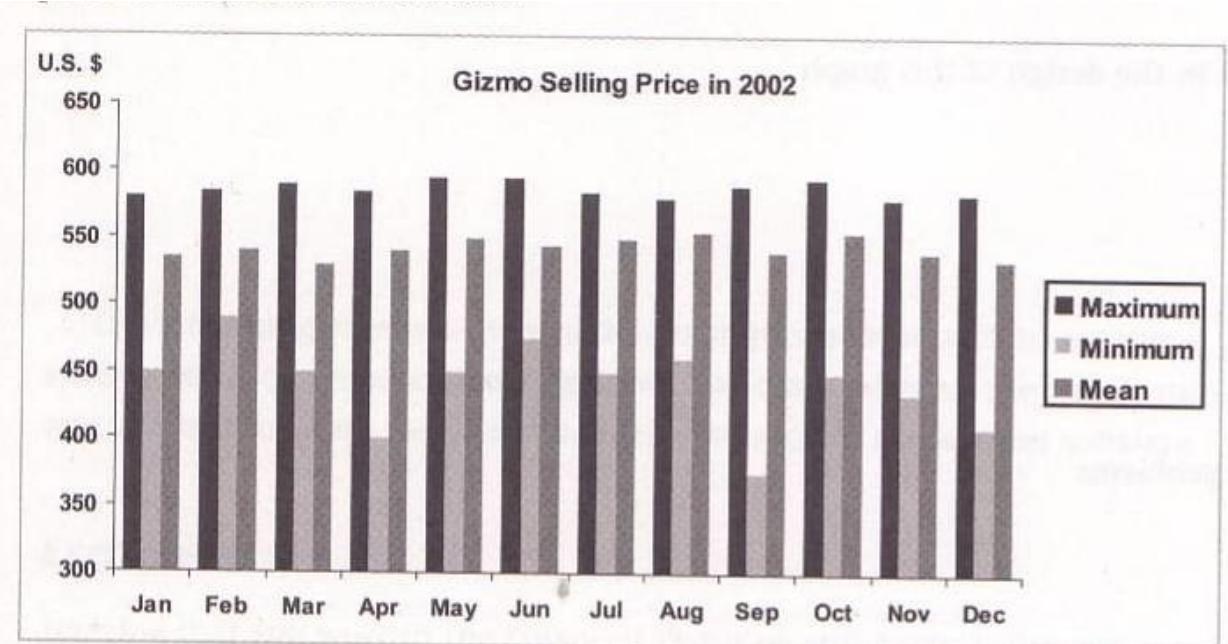
Before



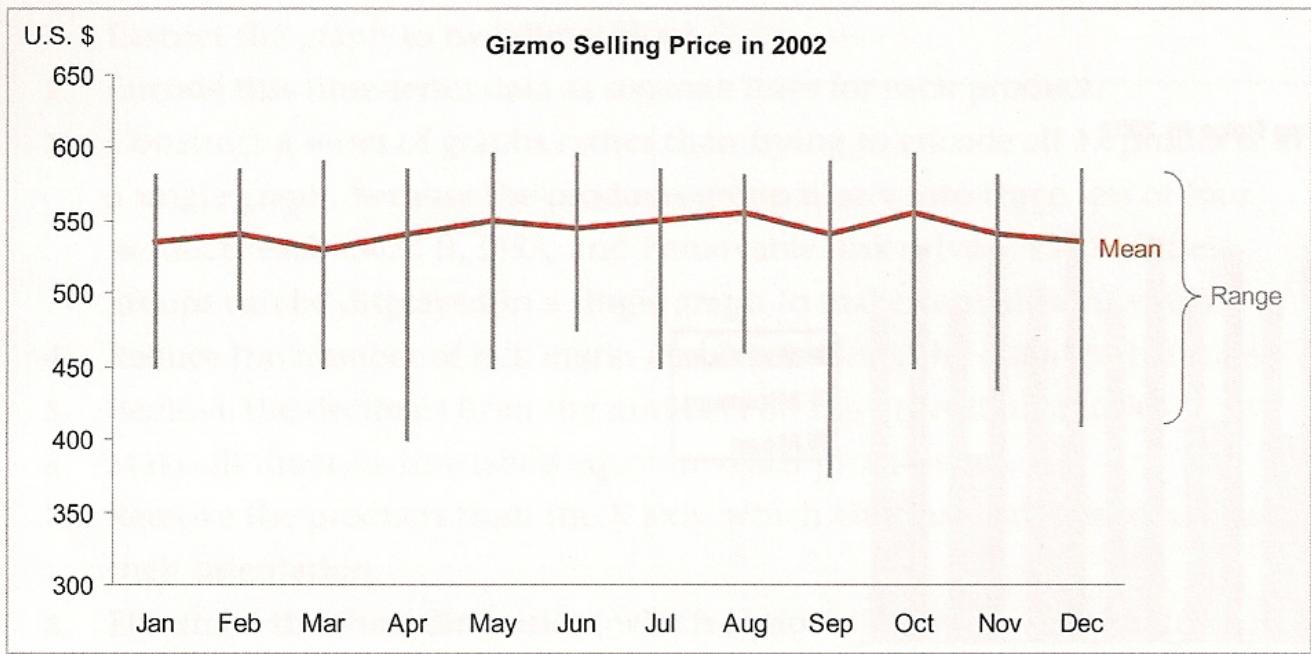
After?



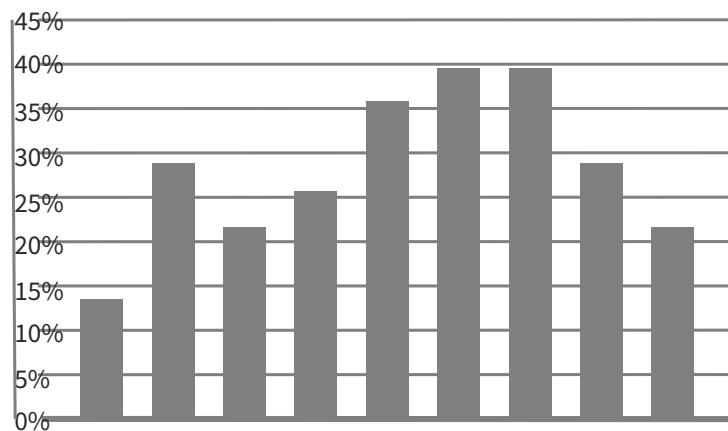
Before



After? less ink is more

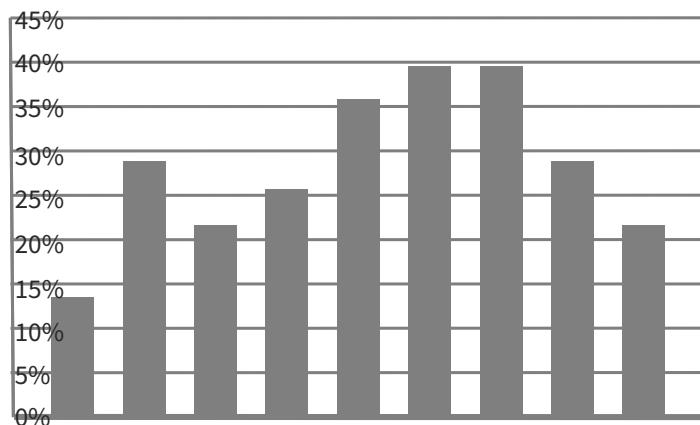


Bar Chart

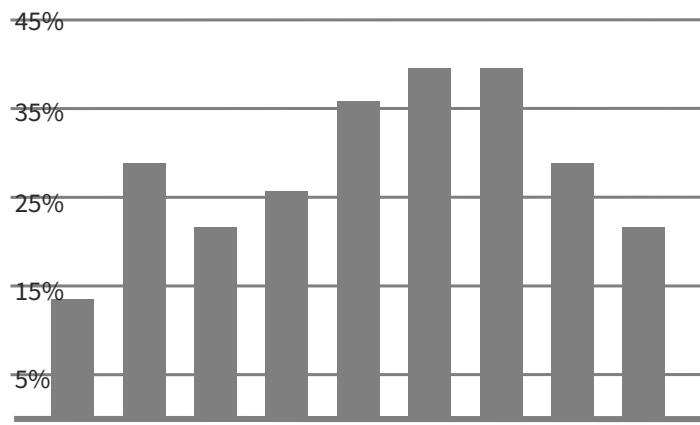


What can we do with this to make it better?

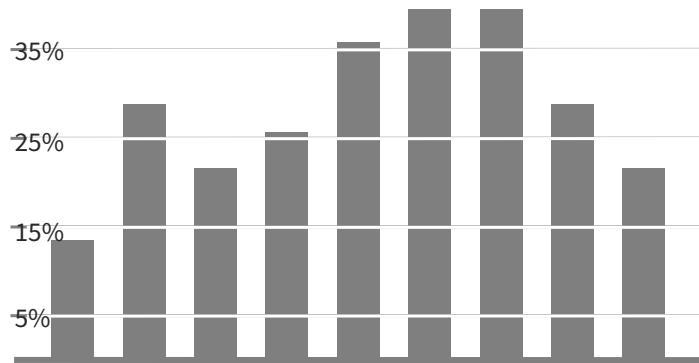
Bar chart



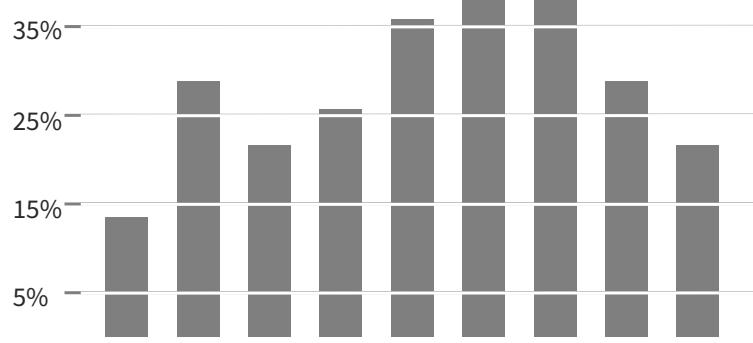
Bar chart



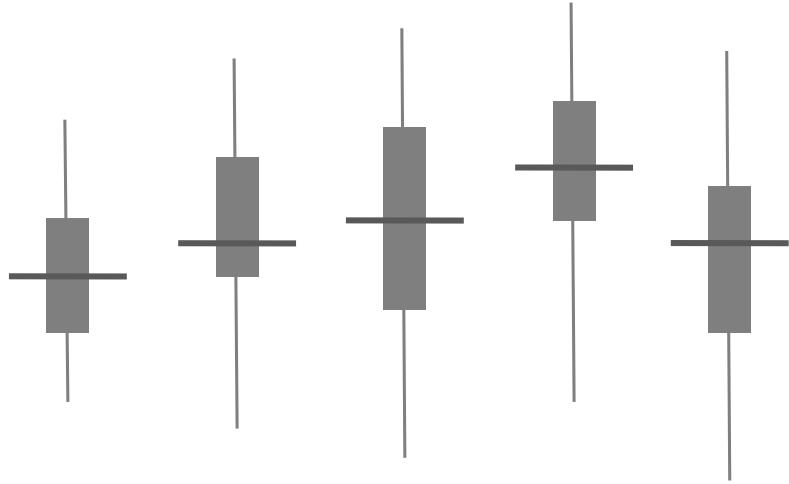
Bar chart



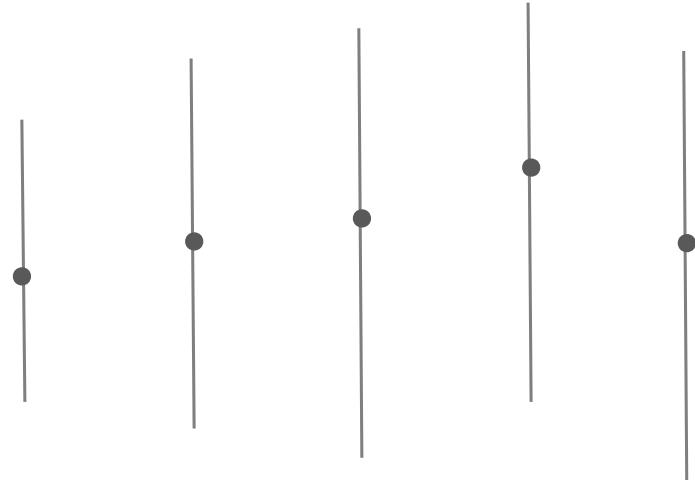
Better Bar Chart



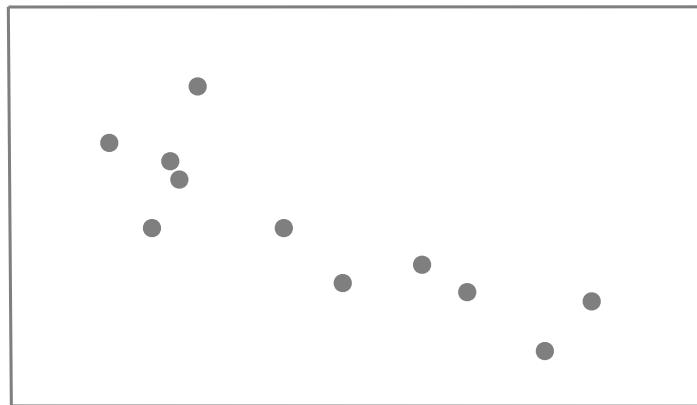
Box Plot



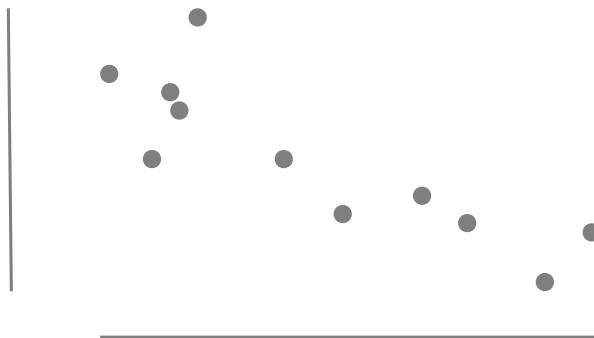
Better Box Plot



Scatterplot



Better Scatterplot



Design Principle

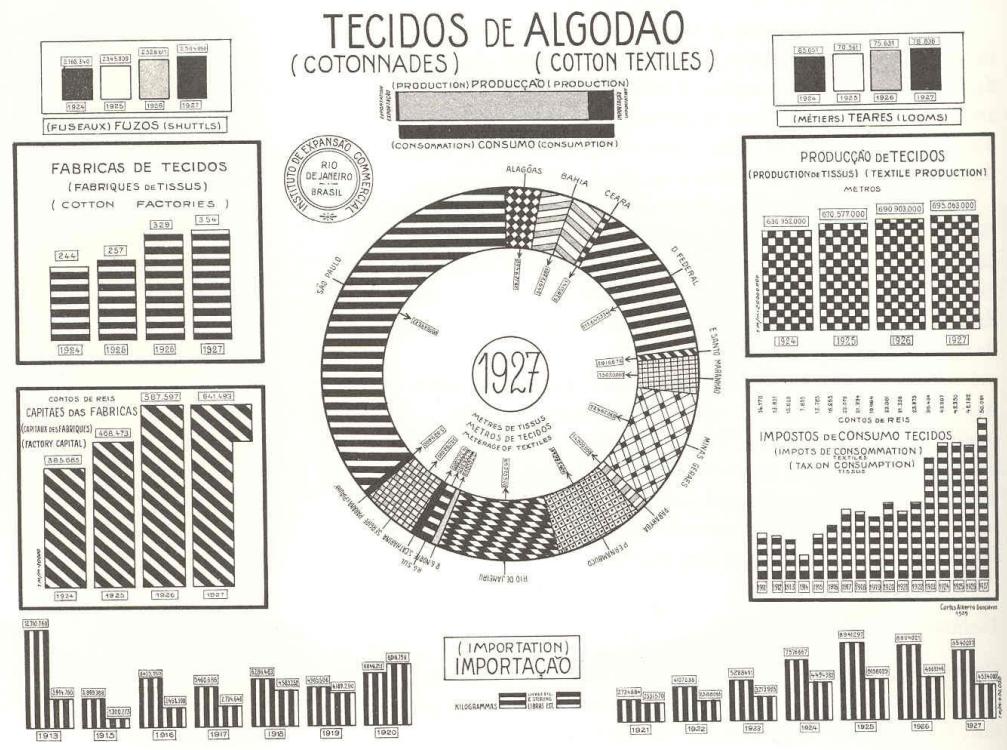
- Avoid **chartjunk**

- Extraneous visual elements that detract from message

some junky examples:

Vol 1, p 108

>_<: Very difficult to know where to focus on



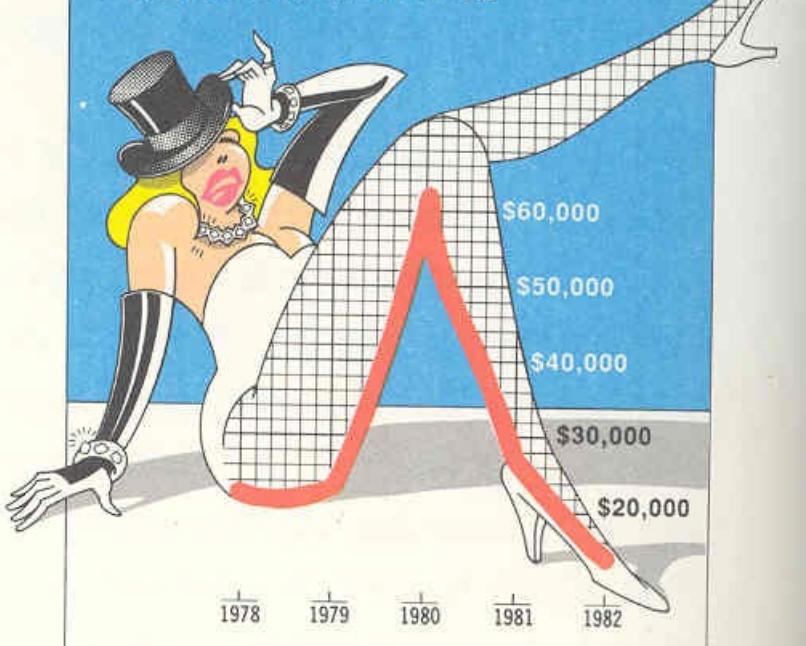
Vol 2, p.34

A classic

Diamonds Were A
Girl's Best Friend

DIAMONDS WERE A GIRL'S BEST FRIEND

Average price of a one-carat D-flawless



Georgia Tech

gatech.edu

USA Today

<http://www.usatoday.com/news/snapshot.htm>

2% = half a cookie??

Advertisement THIS IS MY POST OFFICE. Download My Desktop Post Office™

UNITED STATES POSTAL SERVICE.

USA TODAY Classifieds: [cars.com](#) | [careerbuilder.com](#) | [eHarmony.com](#) | [Marketplace](#) | [Real estate](#)

USA TODAY Snapshot

02/09/2006 - Updated 11:15 PM ET

On Christmas Eve, 48% of adults who say they leave snacks for Santa set out cookies and milk.

Snacks for Santa

Snack	Percentage
Cookies and milk	48%
Cookies	29%
Cookies, milk and carrots	2%

By Anne R. Carey and Keith Carter, USA TODAY
Source: Opinion Research Corp. for Lactaid

Advertisemen

Georgia Tech

gatech.edu

Rethink That?

CHI 2010: Graphs

April 10–15, 2010, Atlanta, GA, USA

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, dam05@mail.usask.ca, cab93@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.

General Terms

Design, Human Factors

INTRODUCTION

Many experts in the area of chart design, such as Edward Tufte, criticize the inclusion of visual embellishment in charts and graphs; their guidelines for good chart design often suggest that the addition of *chart junk*, decorations and other kinds of non-essential imagery, to a chart can make interpretation more difficult and can distract readers from the data [22]. This minimalist perspective advocates plain and simple charts that maximize the proportion of *data-junk* – 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).

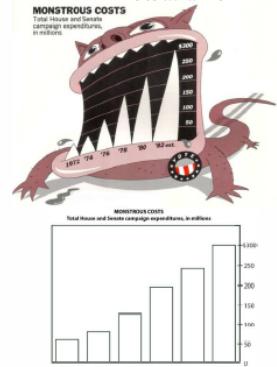


Figure 1. A chart by Holmes [7] (above), and a ‘plain’ version.

Some times embellished charts are just as good on interpretation accuracy as regular charts, and are easier to be recalled weeks later. Whether to choose it depends on who are your audiences and what the audiences gonna to do about your data.

Compared plain charts to “embellished” charts

Found that the embellished charts were just as good on interpretation accuracy and were recalled better weeks later

Participants also preferred the embellished ones

Some caveats:

Very simple data

Very plain plain charts

Each chart/data is different

My take: It's all about purpose

endert@gatech.edu

G Endert

Design Principle

- Utilize **multifunctioning graphical elements**
(macro/micro readings)
- Graphical elements that **convey data information and a design function**

Vol 1, p 140

$0 | 9 = 900 \text{ feet}$

the length of leaves also represents the number of data items fall into that interval of height.

Stem-and-leaf displays:
heights of 218 volcanoes, unit 100 feet.

$19 | 3 = 19,300 \text{ feet}$

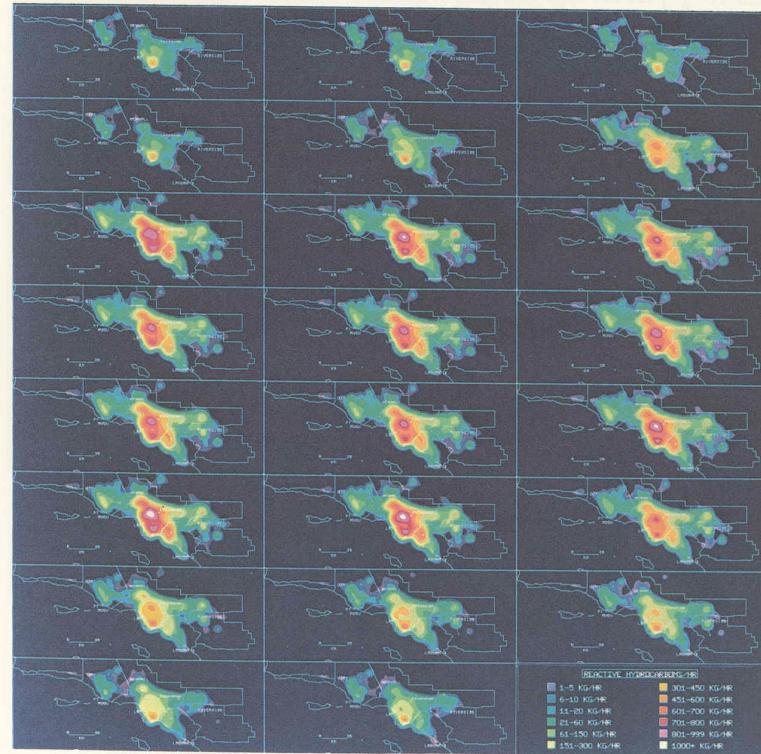
0	98766562
1	97719630
2	69987766544422211009850
3	876655412099551426
4	9998844331929433361107
5	97666666554422210097731
6	898665441077761065
7	98855431100652108073
8	653322122937
9	377655421000493
10	0984433165212
11	4963201631
12	45421164
13	47830
14	00
15	676
16	52
17	92
18	5
19	39730

Design Principle

- Use **small multiples**
 - Repeat visually similar graphical elements nearby rather than spreading far apart

Vol 1, p. 170

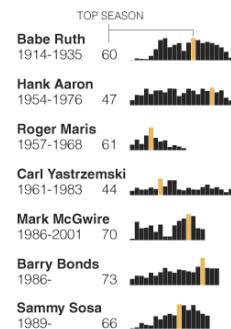
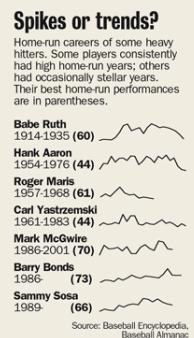
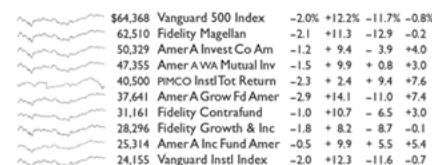
23 hours of
LA air pollution



Georgia Tech

endert@gatech.edu

Sparkline Examples



Georgia Tech

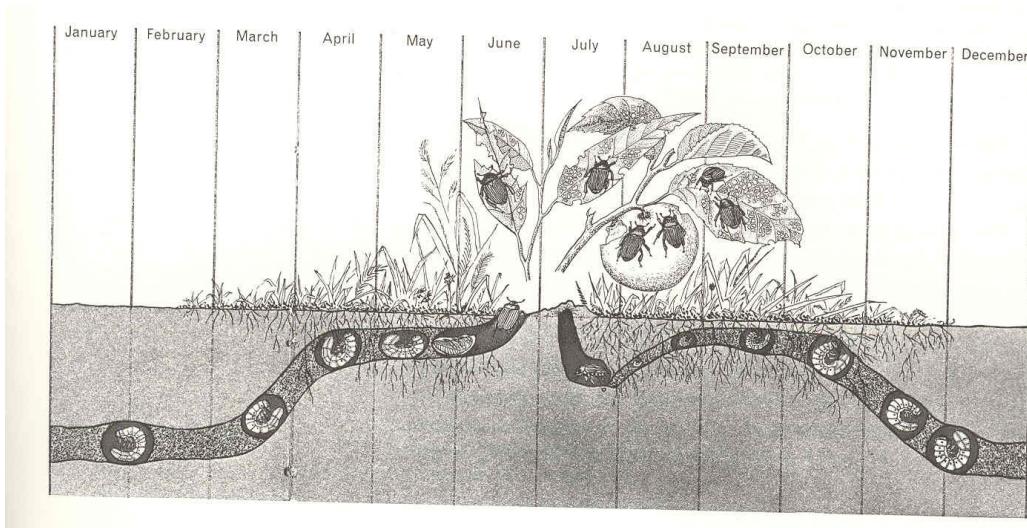
Alex Endert

endert@gatech.edu

Design Principle

- Utilize **narratives** of space and time
 - Tell a story of position and chronology through visual elements

Vol 1, p.43 & Vol 2, p 110



Life of a beetle

L. Hugh Newman

Design Principle

- Content is king
 - Quality, relevance and integrity of the content is fundamental
 - What's the analysis task? Make the visual design reflect that
 - Integrate text, chart, graphic, map into a coherent narrative

Graph and Chart Tips

- Avoid separate legends and keys -- Just have that information in the graphic
- Make grids, labeling, etc., very faint so that they recede into background

Vol 2, p. 54

New Jersey Transit

Train No.	3701	3301	3801	A 67	3803	3201	A3	51	3703	3	3807	3203	A3	61	3809	3	47	3901	3	3903	3	3813	3205	3815	3817	3819	3207	3821	3823	3825	3209	3827	3829	3831
New York, N.Y.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	A.M.	P.M.	P.M.	P.M.			
Newark, N.J. p	12.10	12.40	1.30	3.52	4.50	6.10	6.25	6.35	6.50	7.10	7.30	7.33	7.45	7.50	8.05	8.25	8.40	8.50	9.10	9.40	10.10	10.25	10.40	11.01	11.40	11.50	12.10	12.40	1.10					
North Elizabeth	12.24	12.55	1.44	4.07	5.04	6.24	6.38	6.49	7.04	7.24	7.45	7.47	7.59	8.04	8.19	8.39	8.54	9.04	9.24	9.54	10.24	10.39	10.54	11.24	11.54	12.04	12.24	12.54	1.24					
Elizabeth	12.31	1.03	1.51	.	5.11	6.31	.	6.56	7.11	7.32	.	7.54	.	8.10	.	8.46	9.01	9.11	9.31	10.01	10.31	10.46	11.01	11.31	12.01	12.11	12.31	1.01	1.31					
Linden	12.36	.	1.56	.	5.16	6.35	.	7.01	7.15	7.37	.	7.59	.	8.18	8.31	8.51	9.06	.	9.36	10.06	10.36	.	11.05	11.36	12.06	.	12.36	1.06	1.38					
North Rahway	12.40	1.11	2.00	.	5.20	6.40	.	7.03	7.20	7.42	.	8.03	.	8.24	8.36	8.57	9.10	9.18	9.40	10.10	10.49	10.53	11.10	11.40	12.10	12.18	12.40	1.10	1.40					
Metro Park (Iselin)	12.44	.	2.04	4.26	5.24	.	6.56	7.10	7.25	.	8.04	8.07	8.15	.	8.40	.	9.14	.	9.44	10.14	10.44	.	10.71	11.43	12.14	.	12.44	1.14	1.44					
Metuchen	12.48	.	2.08	.	5.28	.	7.14	7.29	.	8.11	.	8.44	.	9.18	.	9.48	.	10.48	.	11.18	11.48	12.18	.	12.48	1.18	1.48								
Edison	12.51	.	2.11	.	5.35	.	7.05	7.24	7.35	.	8.14	.	8.47	.	9.17	.	9.54	10.25	10.54	.	11.25	11.54	12.25	.	12.54	1.25	1.54							
New Brunswick	12.55	2.15	.	5.35	7.05	7.21	7.35	.	8.25	8.50	.	9.26	9.54	.	10.28	.	11.28	12.28	.	12.28	1.28	1.54												
Jersey Avenue	1.02	.	2.18	.	5.36	.	7.28	.	8.21	.	9.21	.	9.28	.	10.28	.	11.28	12.28	.	12.28	1.28	1.54												
Princeton Jct. S	.	.	2.31	.	5.50	7.19	7.50	.	8.34	8.41	9.05	9.41	.	10.09	10.41	11.09	.	11.41	12.09	12.41	.	1.09	1.41	2.09	.									
Trenton, N.J.	.	.	2.42	4.58	6.03	7.28	.	8.01	8.31	8.44	8.52	9.16	9.52	.	10.15	10.52	11.19	.	11.52	12.19	12.52	.	1.22	1.52	2.20	.								

Before

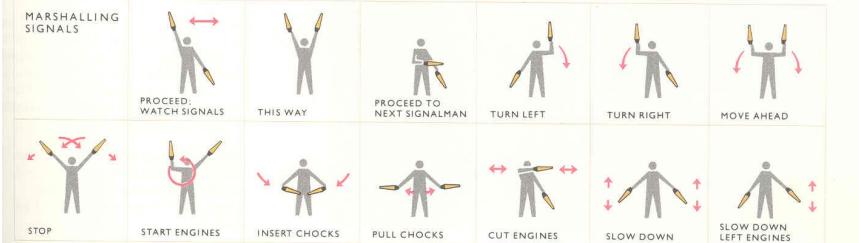
am	•	New York, NY	12.10	12.40	1.30	3.52	4.50	6.10	6.25	6.35	6.50	7.10	7.30	7.33	7.45	7.50	8.04	8.19	8.39	8.54	9.04	9.24	9.40	9.10	10.10	10.25	10.40	11.10	11.24	11.54	
Newark, NJ	•	12.24	12.55	1.44	4.07	5.04	6.24	6.38	6.49	7.04	7.24	7.45	7.47	7.59	8.04	8.19	8.39	8.54	9.04	9.24	9.40	9.10	10.10	10.39	10.54	11.10	11.24	11.54			
North Elizabeth	.	12.31	1.03	1.51	.	5.11	6.31	.	6.56	7.11	7.32	.	7.54	.	8.10	.	8.12	8.26	8.46	9.01	9.11	9.31	10.01	10.31	10.46	11.01	11.31	12.01			
Elizabeth	.	12.31	1.03	1.51	.	5.11	6.31	.	6.56	7.11	7.32	.	7.54	.	8.18	8.31	8.51	9.06	.	9.36	10.06	10.36	.	11.06	11.36	12.06	.				
Linden	.	12.36	.	1.56	.	5.16	6.36	.	7.01	7.15	7.37	.	7.59	.	8.20	8.33	8.54	.	9.10	9.18	9.40	10.10	10.40	10.53	11.10	11.40	12.10				
North Rahway	.	Rahway	12.40	1.11	2.00	.	5.20	6.40	.	7.06	7.20	7.42	.	8.03	.	8.24	8.36	8.57	9.10	9.18	9.40	10.10	10.40	10.53	11.10	11.40	12.10				
Metro Park (Iselin)	.	12.44	.	2.04	4.26	5.24	.	6.56	7.10	7.25	.	8.04	8.07	8.15	.	8.40	.	9.14	.	9.44	10.14	10.44	.	11.14	11.44	12.14	.				
Metuchen	.	12.48	.	2.08	.	5.28	.	7.14	7.29	.	8.11	.	8.44	.	9.18	.	9.48	10.18	10.48	.	11.18	11.48	12.18	.	12.21	.					
Edison	.	12.51	.	2.11	.	5.35	.	7.17	7.32	.	8.14	.	8.47	.	9.21	.	10.21	.	11.21	.	12.21	.									
New Brunswick	.	12.55	2.15	.	5.35	7.05	7.21	7.35	.	8.18	8.25	8.50	9.25	.	9.54	10.25	10.54	.	11.25	11.54	12.25	.									
Jersey Avenue	.	1.02	2.18	.	5.36	7.08	7.28	.	8.21	.	8.21	.	9.28	.	10.28	.	11.28	12.28	.	12.28	1.28	1.54									
Princeton JunctionS	.	2.31	.	5.50	7.19	7.50	.	8.34	8.41	9.05	9.41	.	10.09	10.41	11.09	.	11.41	12.09	12.41	.											
Trenton, NJ	.	2.42	4.58	6.03	7.28	8.01	8.31	8.44	8.52	9.16	9.52	.	10.19	10.52	11.19	.	11.52	12.19	12.52	.											

After

Vol 2, p. 63



Before
After



In summary, graphical Displays Should

- Show the data
- Induce the viewer to think about substance rather than about methodology, graphic design the technology of graphic production, or something else
- Avoid distorting what the data have to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several levels of detail, from a broad overview to the fine structure
- Serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- Be closely integrated with statistical and verbal descriptions of a data set

Sources Used

- E. Tufte, The Visual Display of Quantitative Information
- E. Tufte, Envisioning Information
- E. Tufte, Visual Explanations