Design Principles

As told by Edward Tufte, Donna Wong, Stephen Few, Jim Foley

CS 4460 - Information Visualization

Jim Foley

Last update November 2016

Sources

- Tufte, *The Visual Display of Quantitative Information*
- Tufte, *Envisioning Information*
- Tufte, *Visual Explanations*
- Wong, *The WSJ Guide to Information Graphics*
  - Wong is a student of Tufte's
- Few, *Now you see it* + web site
- Foley

Graphical Excellence is...

- The presentation of interesting data is a matter of *substance*, of *data*, and of *design*.
    - *Substance* - having a message
    - *Data* - having right data for the message
    - *Design* - having right design
        Or at least a good design
        "Right design" is somewhat subjective
- Communication of complex ideas with *clarity, precision* and *efficiency*.

Graphical Excellence

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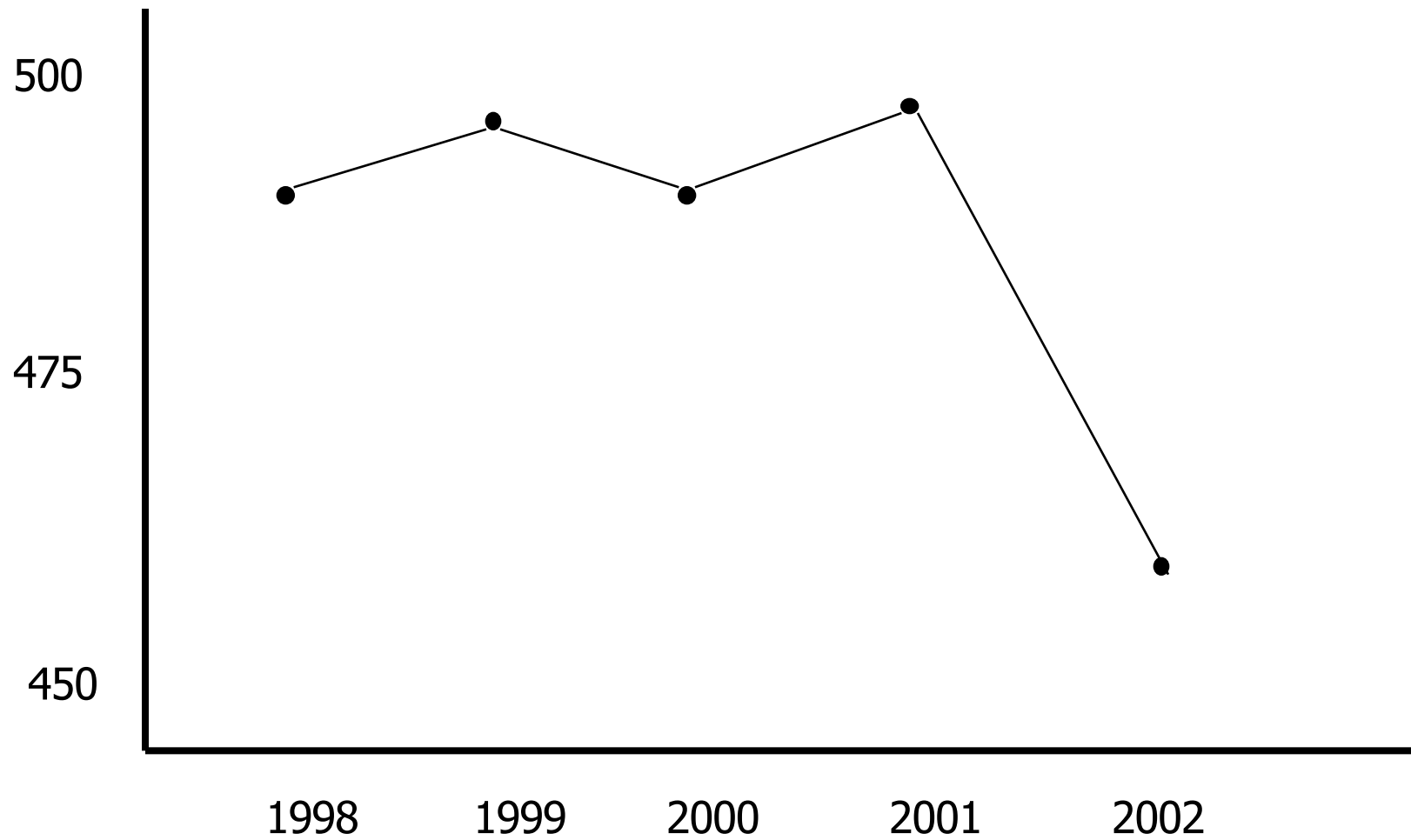 - Design aesthetics

Let's look at each of these

1. Graphical Integrity

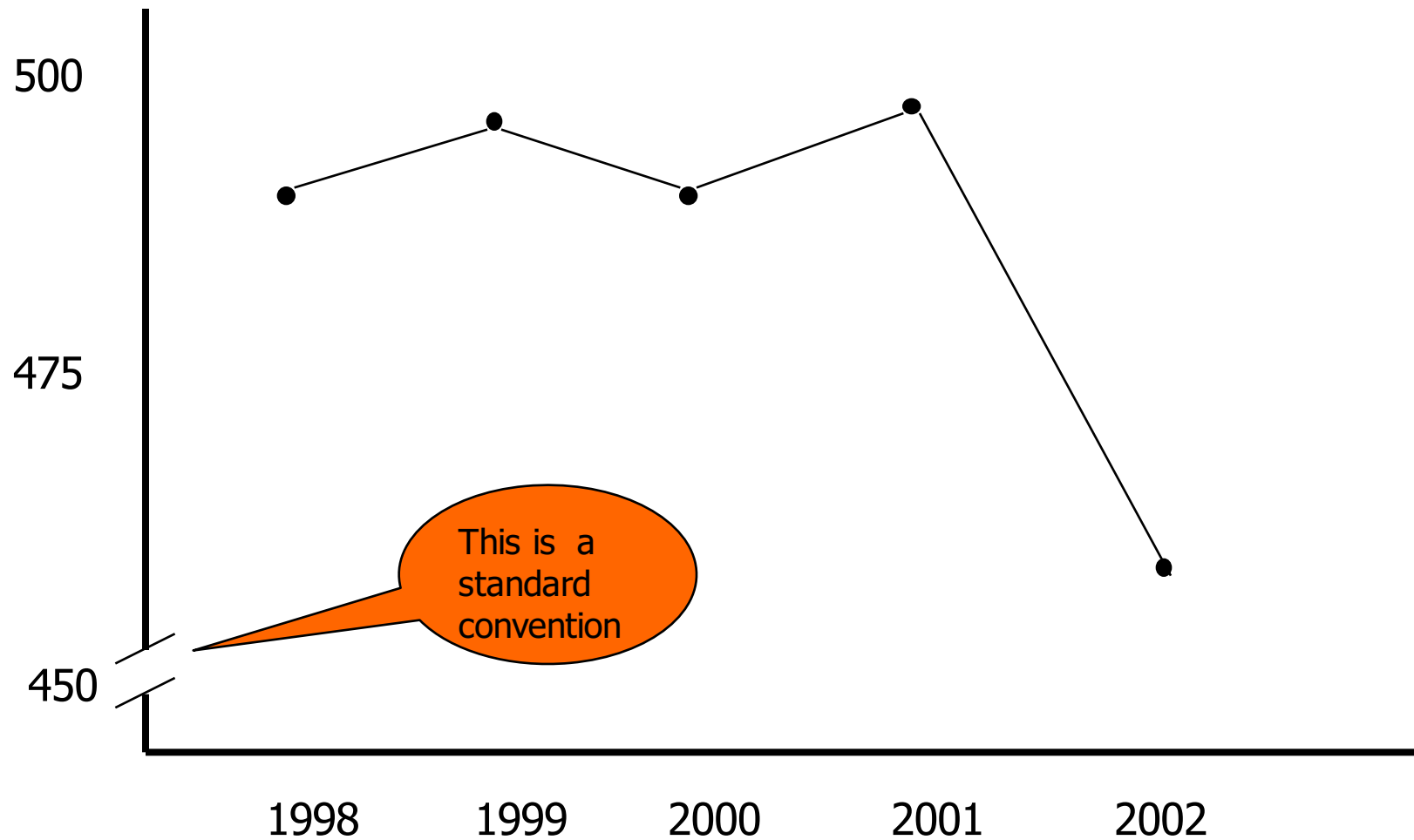- Your graphic should **tell the truth** about your data

- Do **NOT** lie
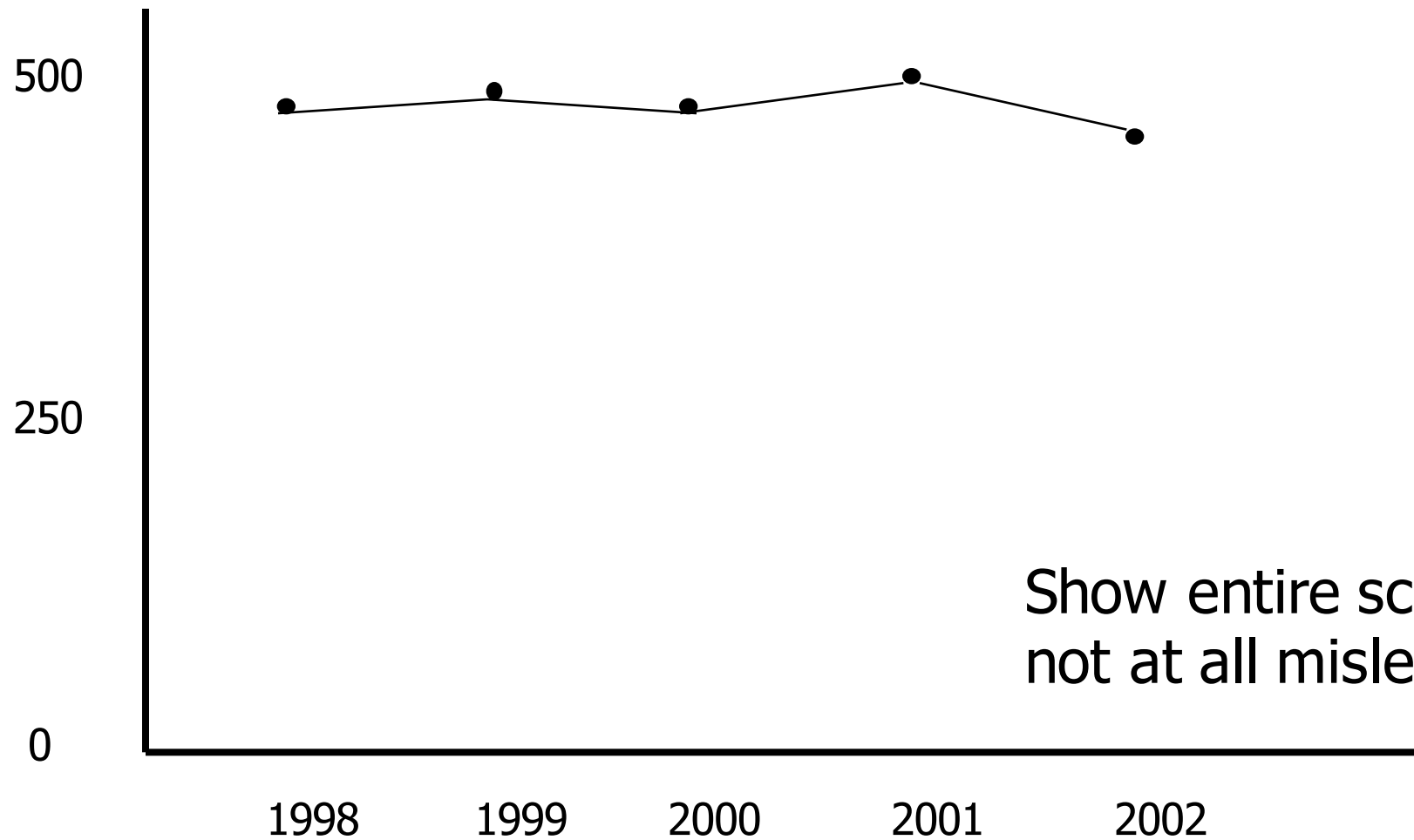
Example

Stock market crash?

500

475

450

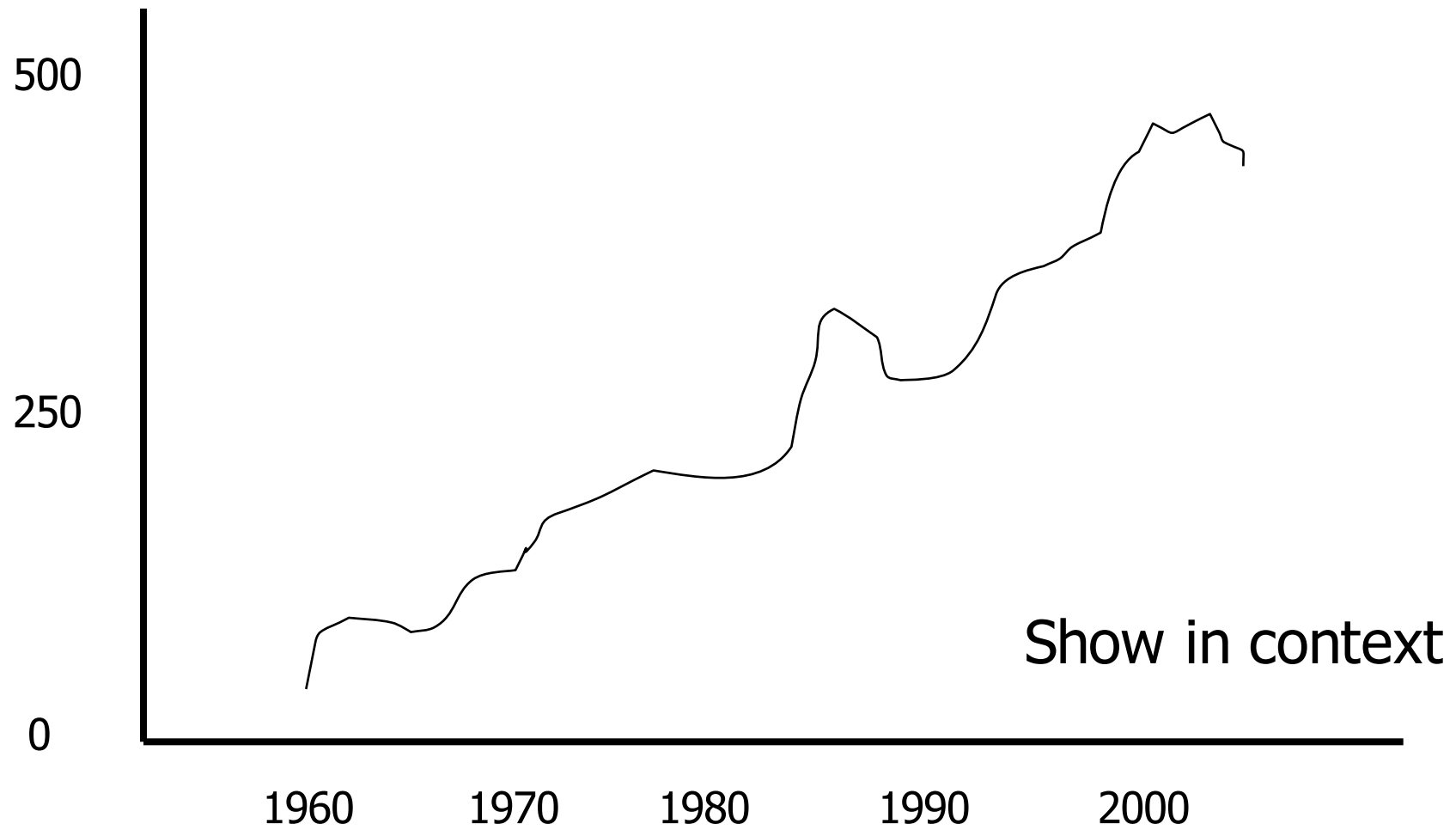1998   1999   2000   2001   2002

Example

Stock market crash?

Example

500

250

0

Show entire scale – not at all misleading

1998   1999   2000   2001   2002

Example

Show in context

Example – what is wrong here?

Varilux S Series™ Technical Paper | 2012 | SynchronEyes Technology™ | 4
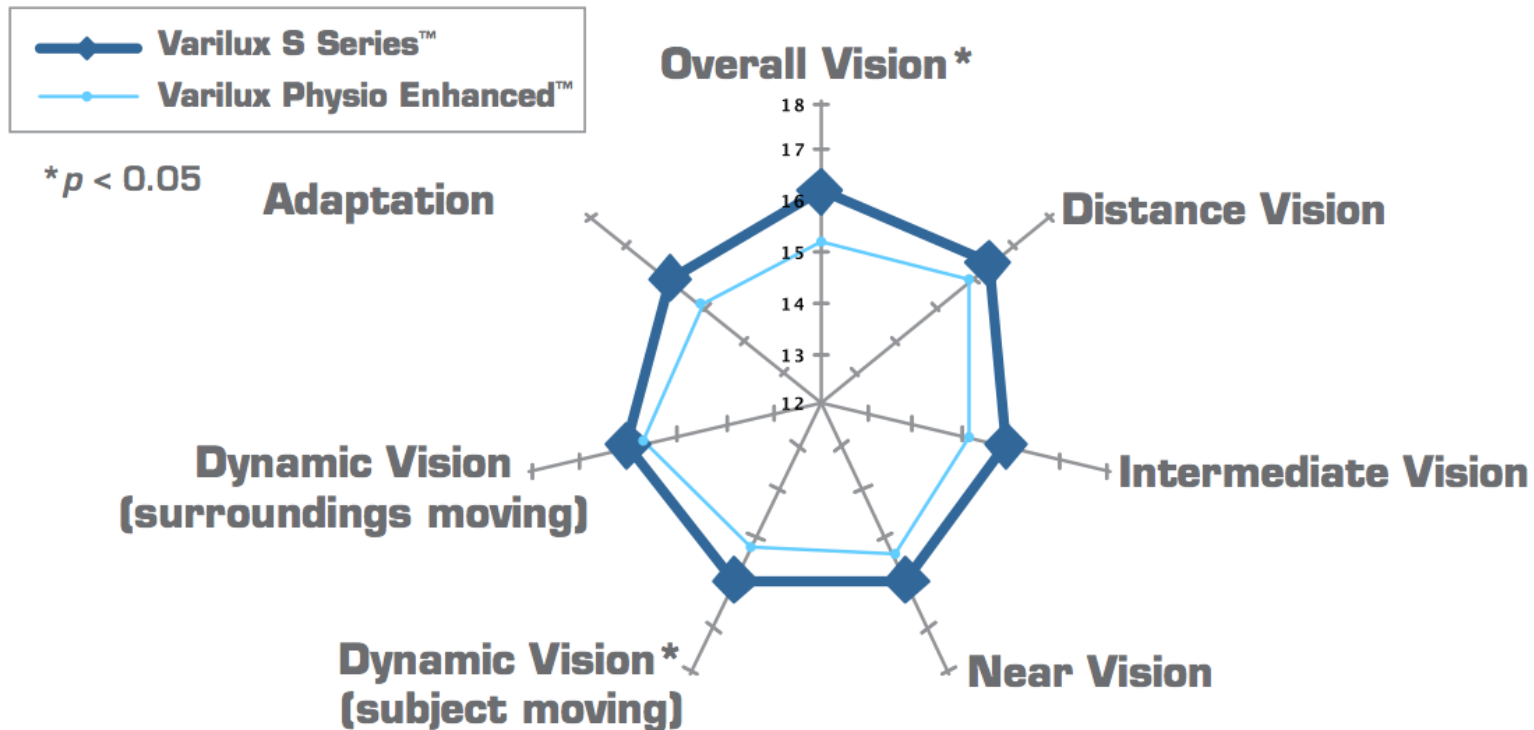
FIGURE 6 In a double-masked crossover study of 97 experienced progressive lens wearers, Varilux S Series™ lenses were preferred over highly regarded Varilux Physio Enhanced™ lenses in every parameter tested.
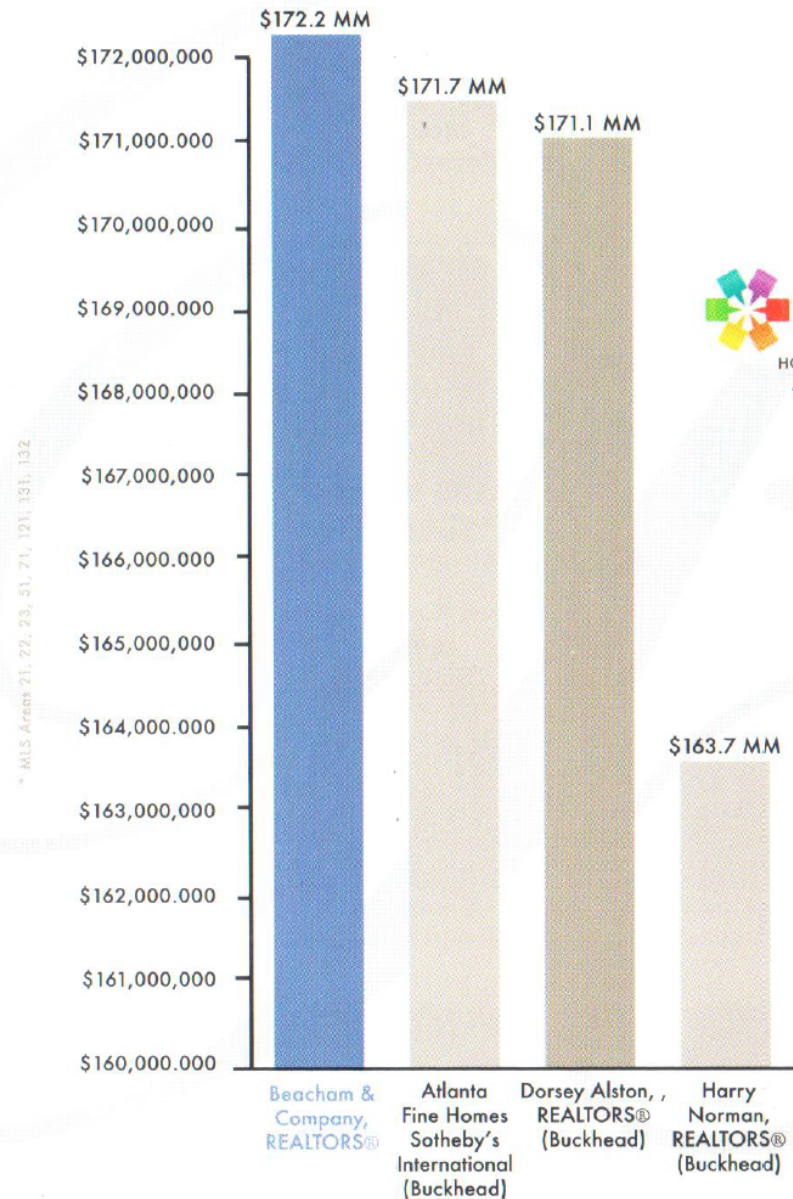
Telling the Truth?

Chart Integrity

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

Where is Origin; Multiple Scales!
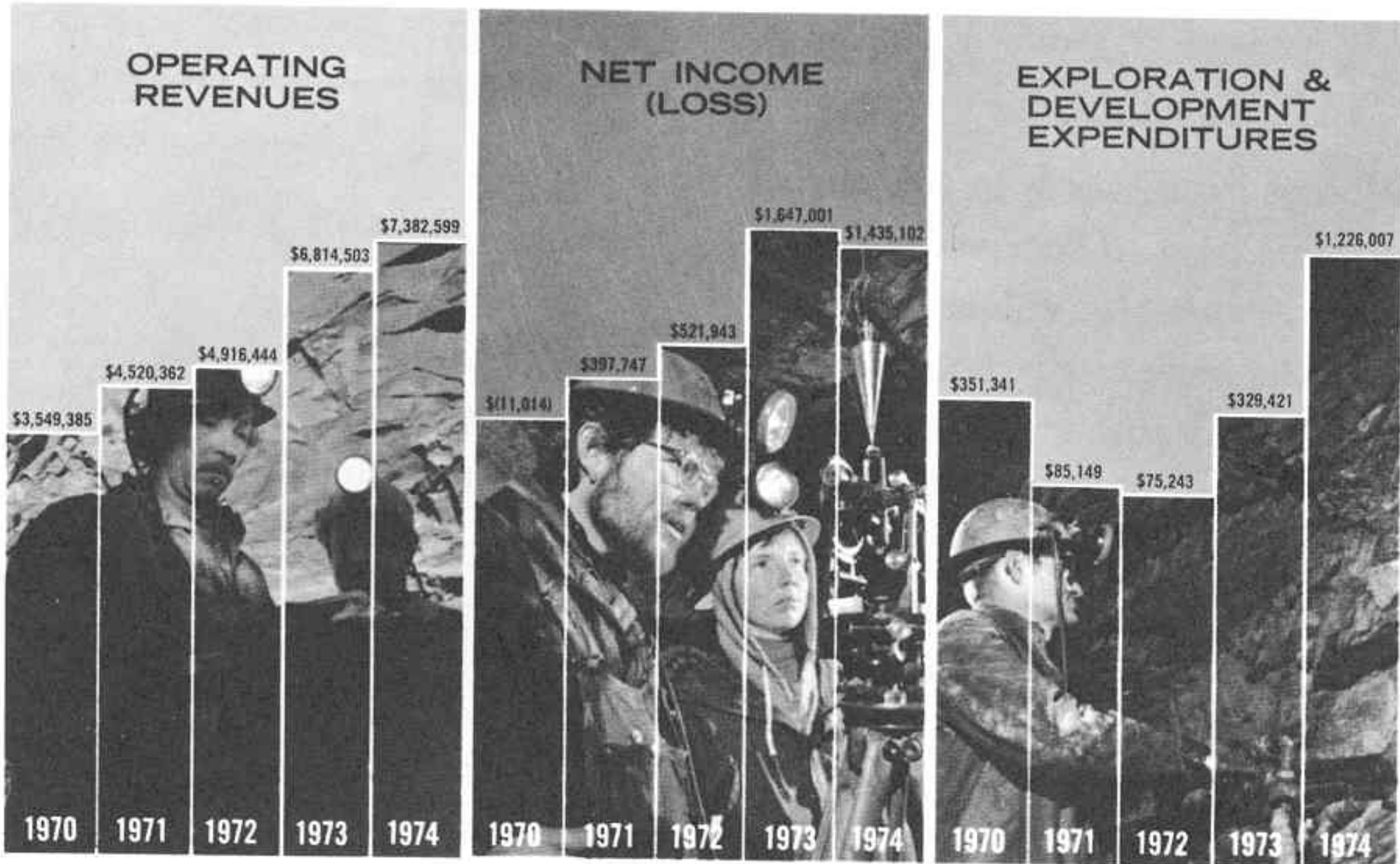
Multiple Scales
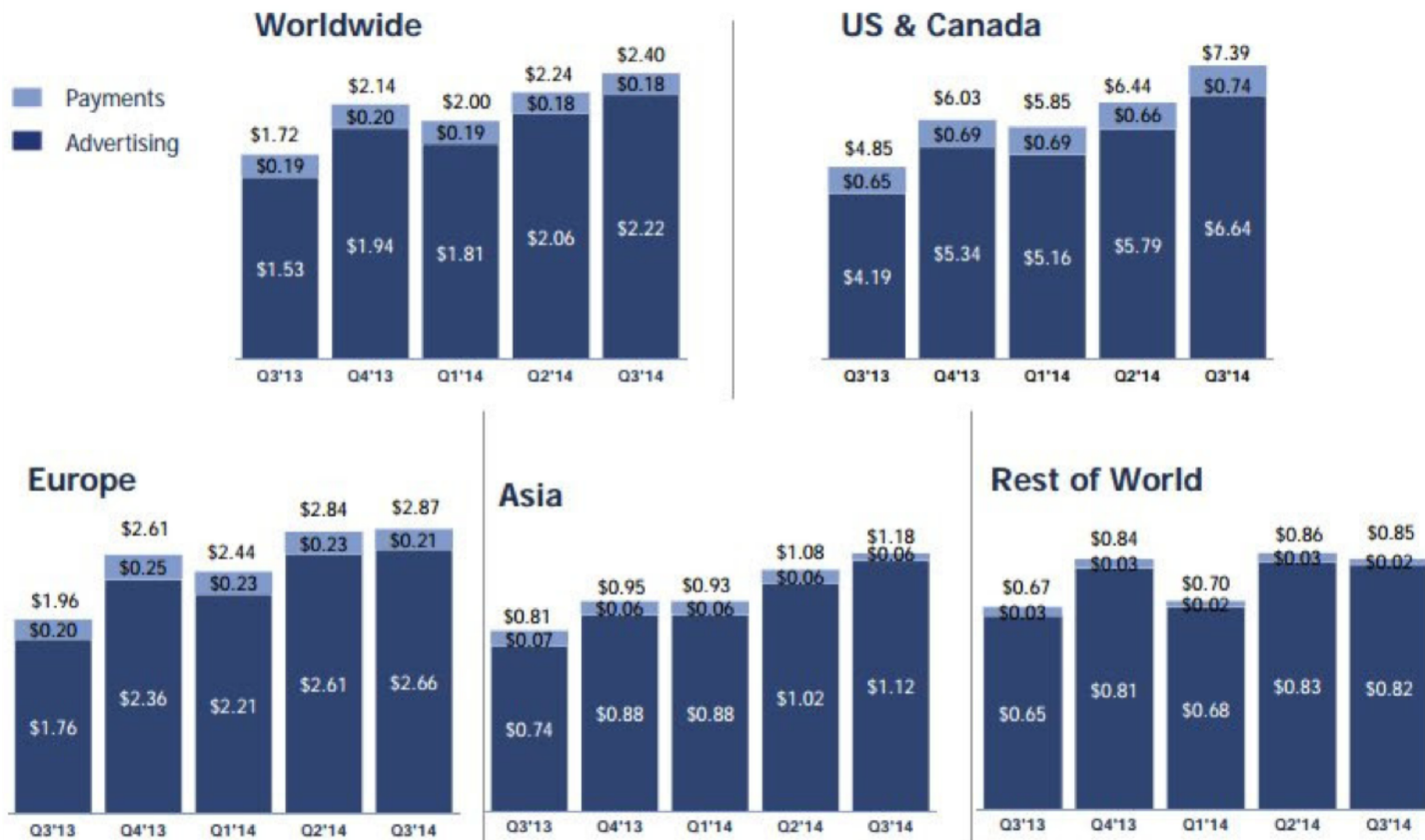
Where's the Context?

325 • Before stricter enforcement

Connecticut Traffic Deaths, Before (1955) and After (1956) Stricter Enforcement by the Police Against Cars Exceeding Speed limit

300
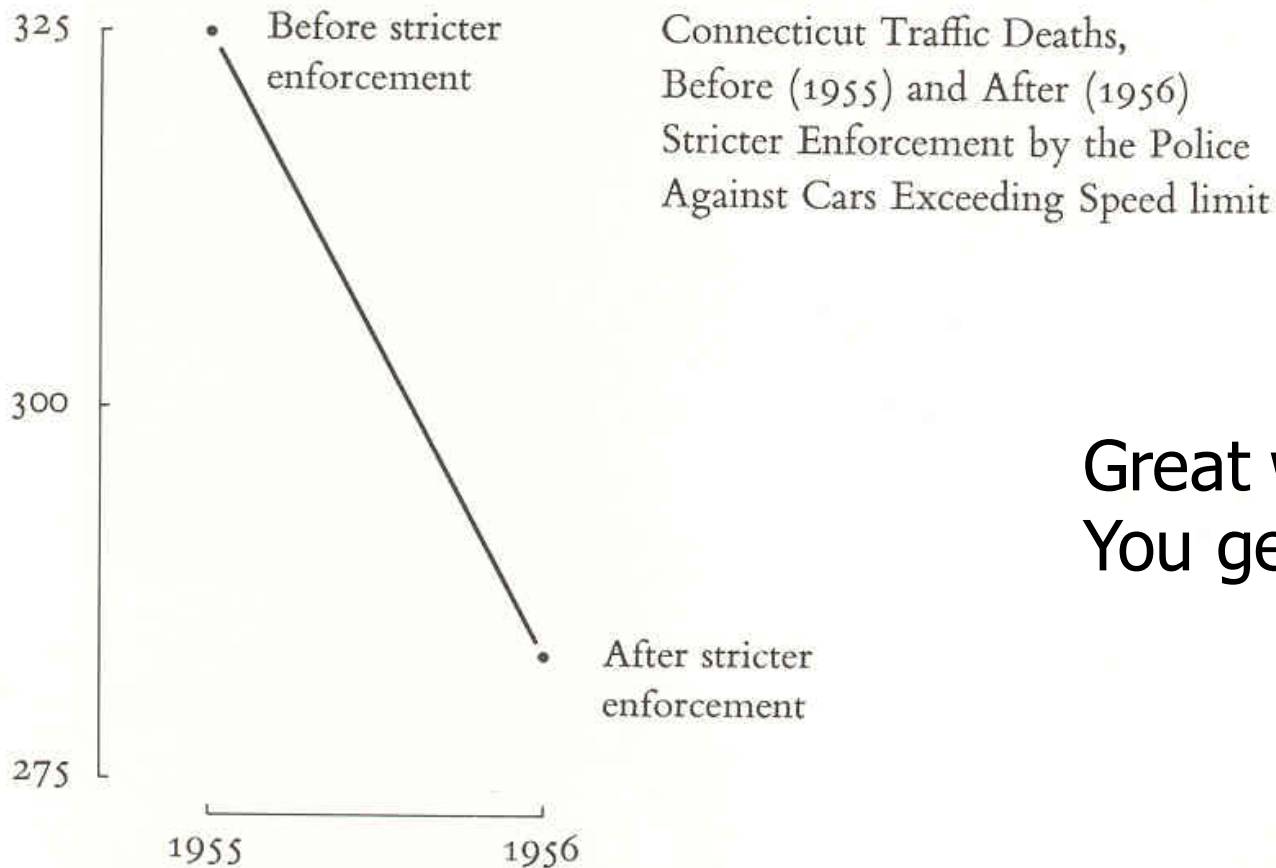
Great work, Connecticut
You get a prize !

After stricter enforcement

275

1955   1956

Tufte Vol 1, p. 74

Here's the Context!

A few more data points add immensely to the account:

Connecticut Traffic Deaths,
1951–1959
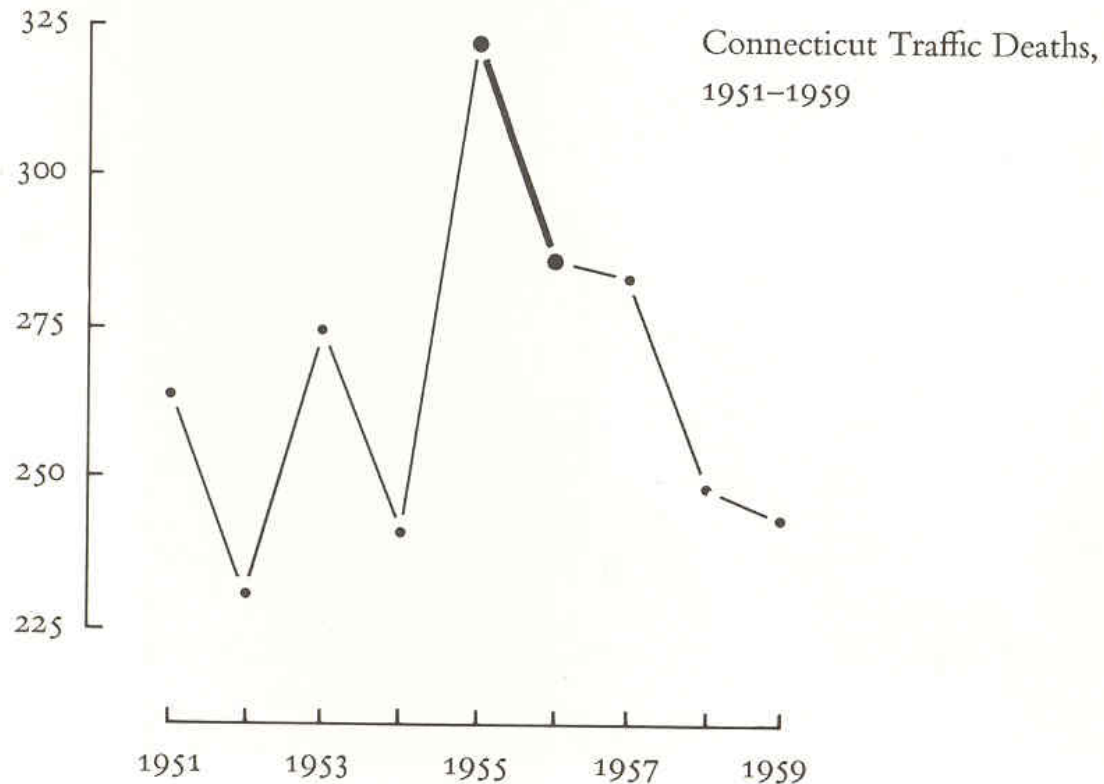
Oops, not so fast!!

Show the context!

Tufte Vol 1, p. 74

Watch Size Coding

- Height/width vs. area vs. volume

Height or Volume?
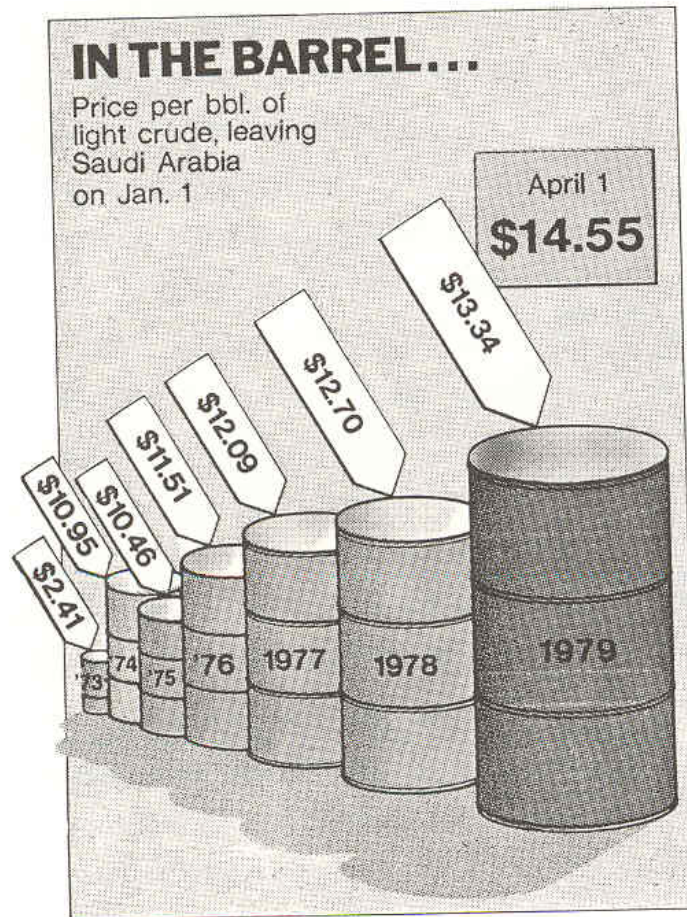
What is encoding price per barrel?
  Height?
  Volume?

Tufte Vol 1, p. 62

Measuring Misrepresentation

- Visual attribute value should be directly proportional to data attribute value

$$\text{Lie factor} = \frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}}$$
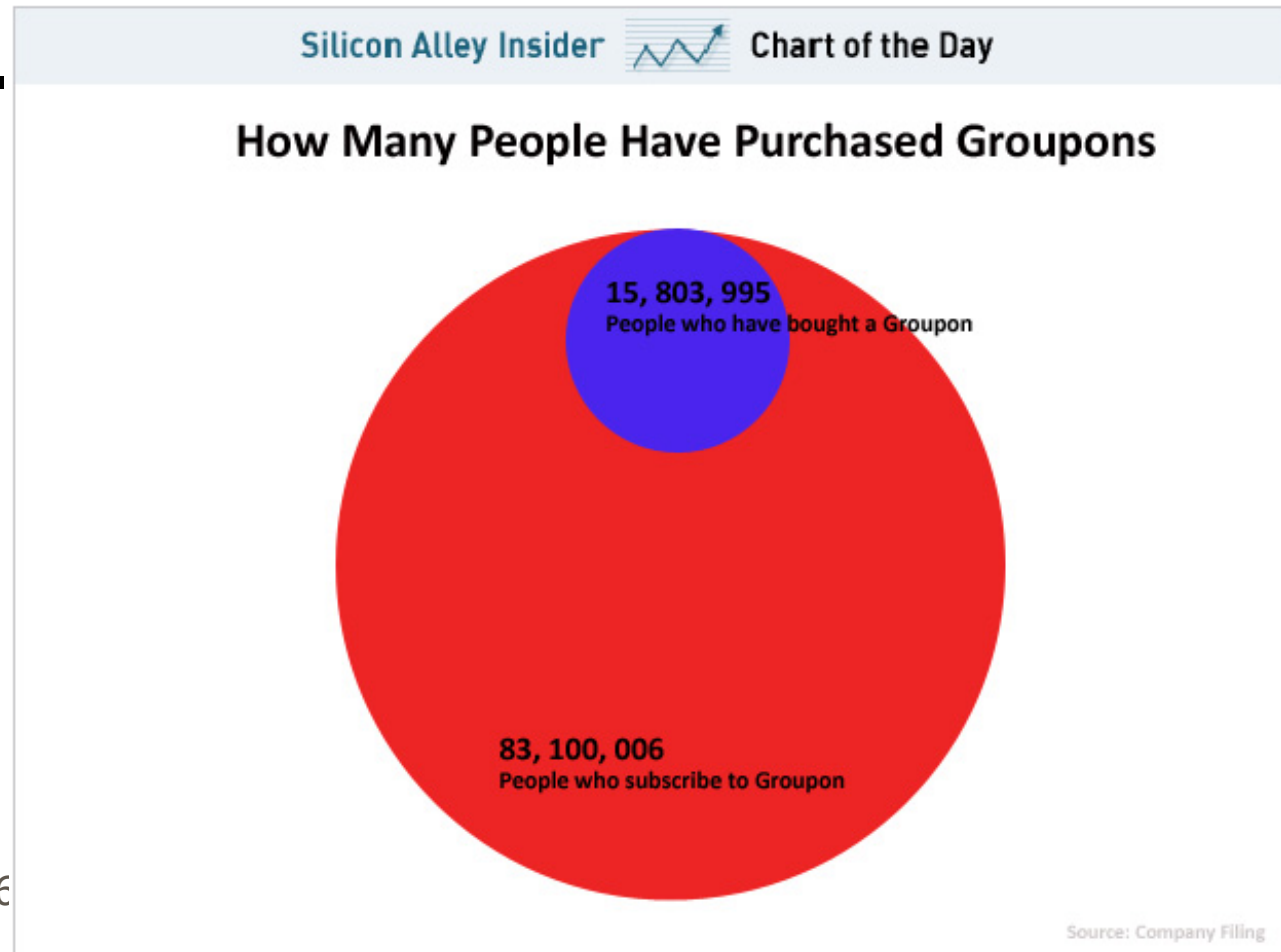
If volume encodes price, then lie factor is 4280/454 = 9.4

Lie Factor

- Graphical area ratio ~9.7

- Data ratio ~5.

- Lie factor 9.7/5.3 = 1.8

Silicon Alley Insider — Chart of the Day

How Many People Have Purchased Groupons

15, 803, 995
People who have bought a Groupon

83, 100, 006
People who subscribe to Groupon

Source: Company Filing

CS 446

Lie Factor in 2011 State of the Union

AMERICA'S ECONOMY
2010 GROSS DOMESTIC PRODUCT

United States $14.6 TRILLION

China $5.7 TRILLION

Japan $5.3 TRILLION

Germany $3.3 TRILLION

France $2.5 TRILLION

US Economy
~6 x France
US Area
~36 x
France Area
Lie Factor =
36/6 = 6

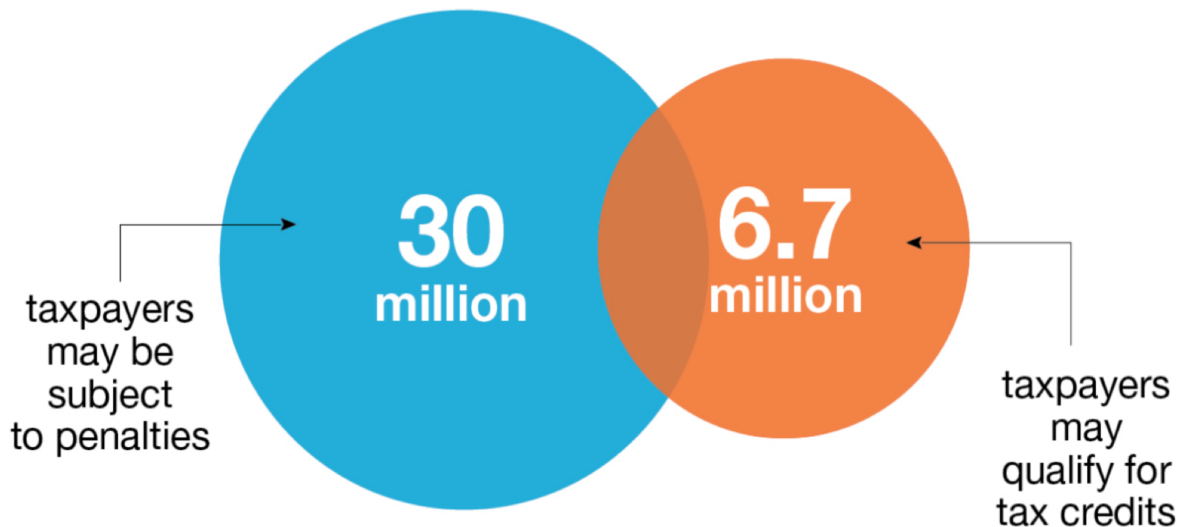http://flowingdata.com/category/statistics/mistaken-data/

Lie Factor

MIGRAINE SURGERY

Surgical migraine treatments focus on decompressing trigger area nerves that cause migraines. Area nerves can be in the forehead, nasal region, area behind the eyes or the back of the head or neck.

A study in the Journal of Plastic and Reconstructive Surgery followed patients 5 years after treatment and found:

29% HAD COMPLETE ELIMINATION OF MIGRAINES

88% maintained overall improvement in frequency, intensity or duration of migraines

59% had significant decrease in migraines

Research shows migraine surgery is a viable treatment route for patients suffering from chronic migraines who have exhausted other medical and non-medical options.

30 million taxpayers may be subject to penalties

6.7 million taxpayers may qualify for tax credits
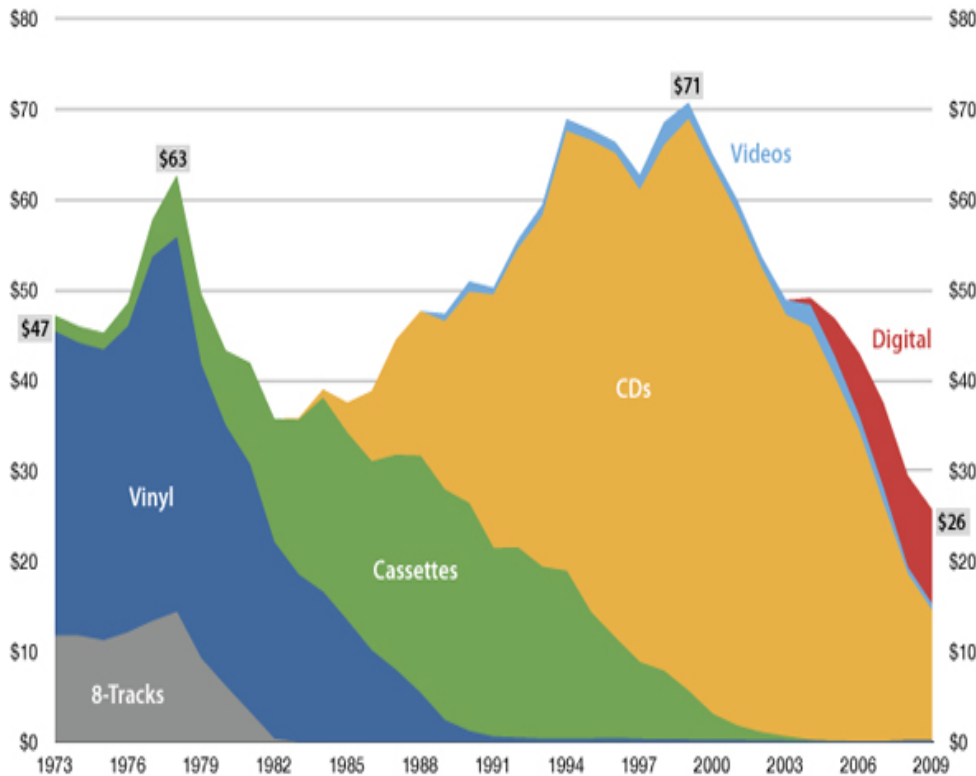
22

Adjust for Inflation and Population

- Financial data can be misleading if not adjusted for inflation

- Per-capita data can show some trends better than totals

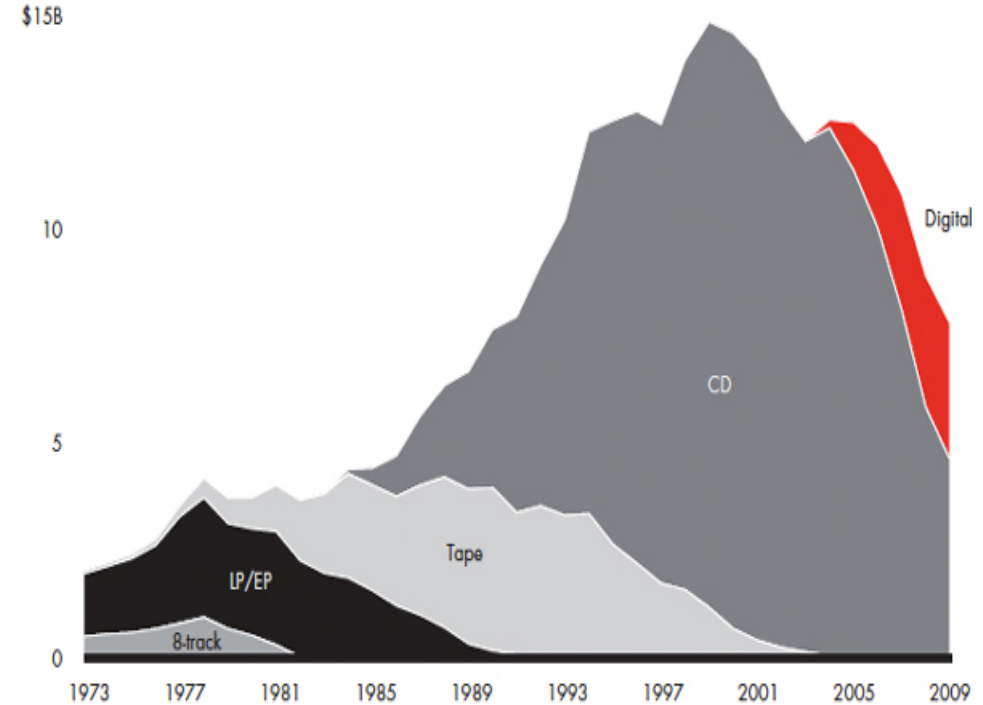- (Not from Tufte)

Adjusted and Unadjusted Data

US Recorded Music Revenue - 2011 Dollars Per Capita

Data: Recording Industry Association of America
Chart/Analysis: Michael DeGusta

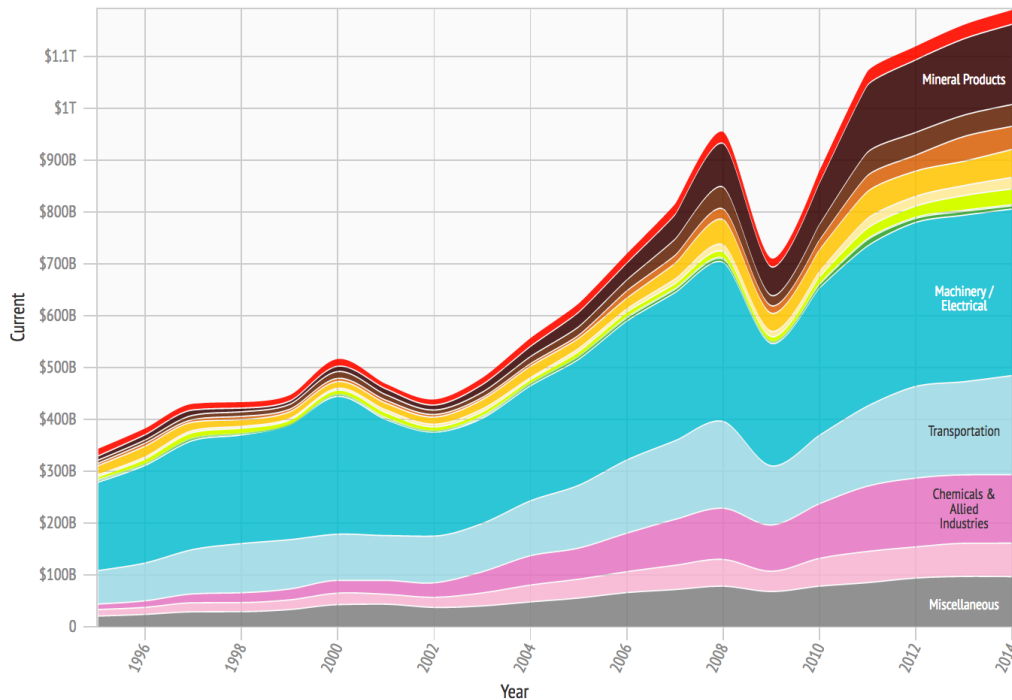Global music industry turnover (1973–2009)

Sources: RIAA year-end shipment statistics; Bain analysis

Adjusted and Unadjusted Data

25

Color-coded Maps Can Mislead

Map
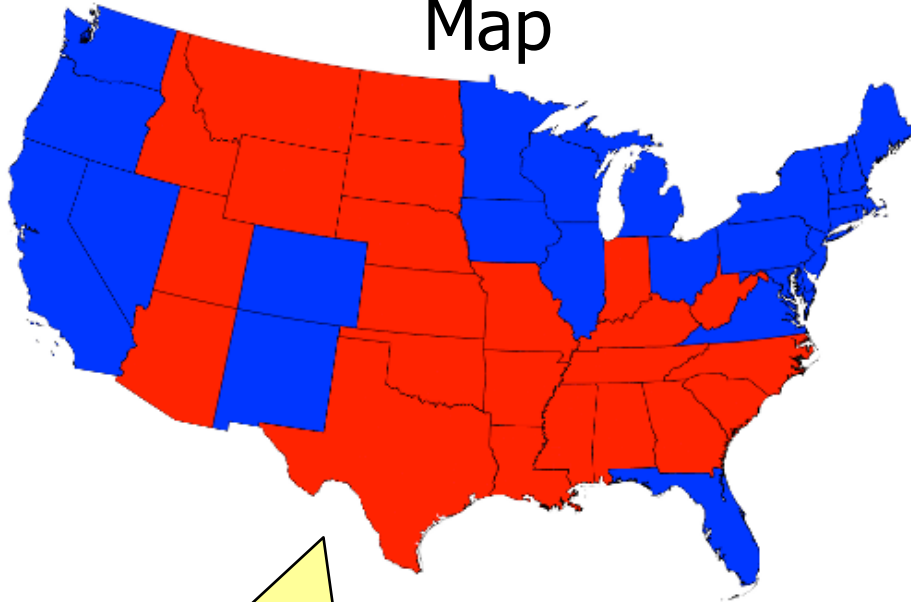
2012 Presidential Election with and without adjustment for each state's Electoral College votes

Cartogram

More red than blue: Republican victory?

More blue than red; Democratic victory.

CS 4460

Variation on Cartogram

States Where Insured Could Renew Plans Before Change by Obama

Legend:
- ■ Allowed early renewals in 2013
- ■ Did not allow
- ■ Information was not available

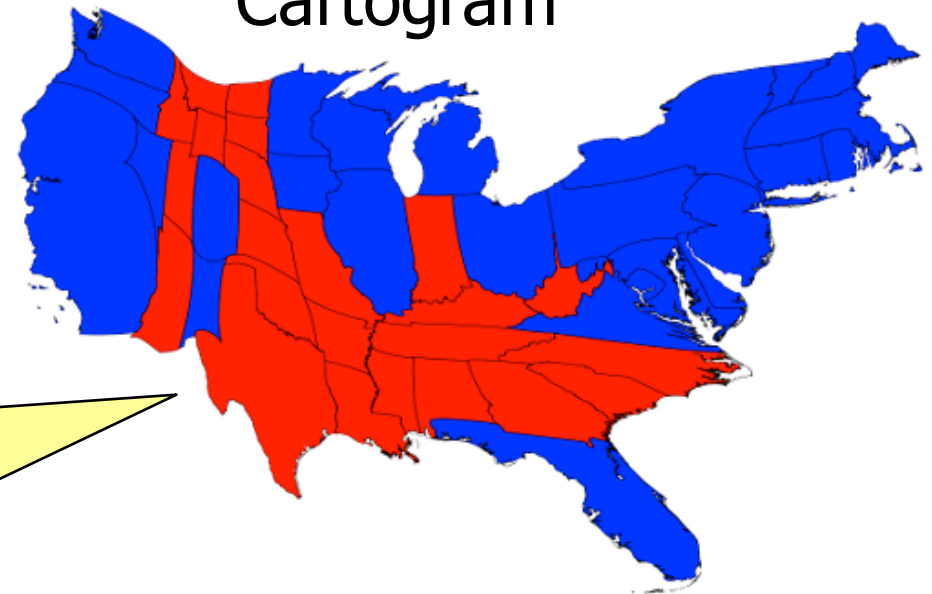Squares are sized according to the number of policies in each state's individual insurance market in 2012.

http://www.nytimes.com/interactive/2013/11/20/us/which-states-will-allow-old-health-policies-to-be-renewed.html?ref=politics
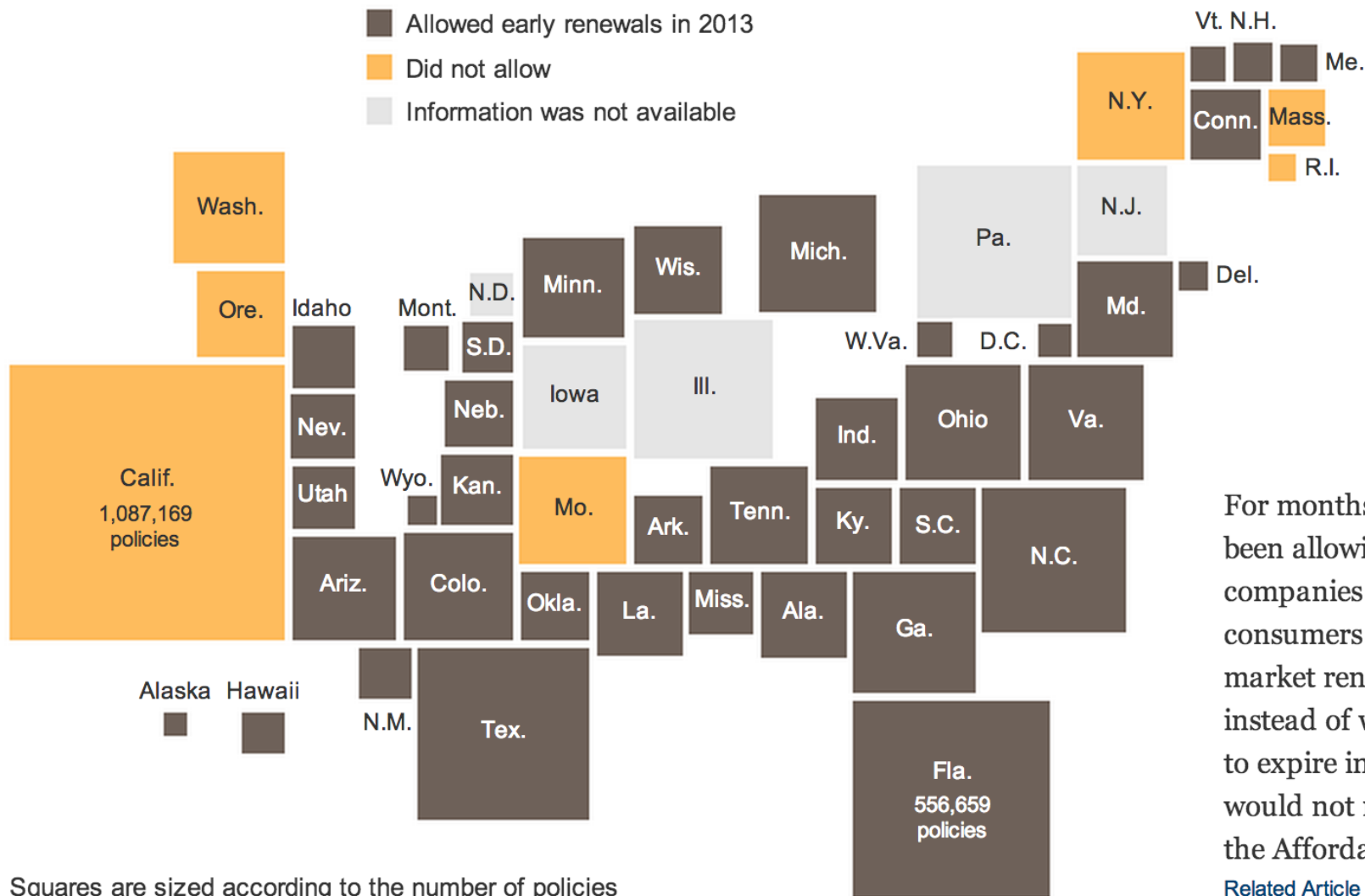
For months, many states have been allowing insurance companies the option of letting consumers in the individual market renew their policies early, instead of waiting for the policies to expire in 2014, when they would not meet the standards of the Affordable Care Act.

Related Article »

Per-capita Data Often More Useful

- Which data is more useful for comparison purposes (made-up data):
  - Alaska and California both had 100 violent crimes last year
  - Alaska had 100 violent crimes per million inhabitants and Texas had 10 violent crimes per million inhabitants

Which Map Best

- For count of states of one color?

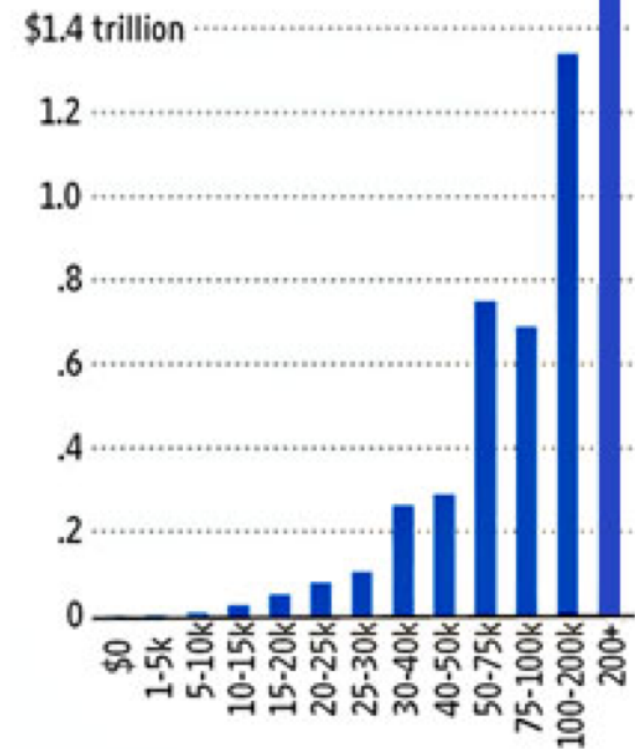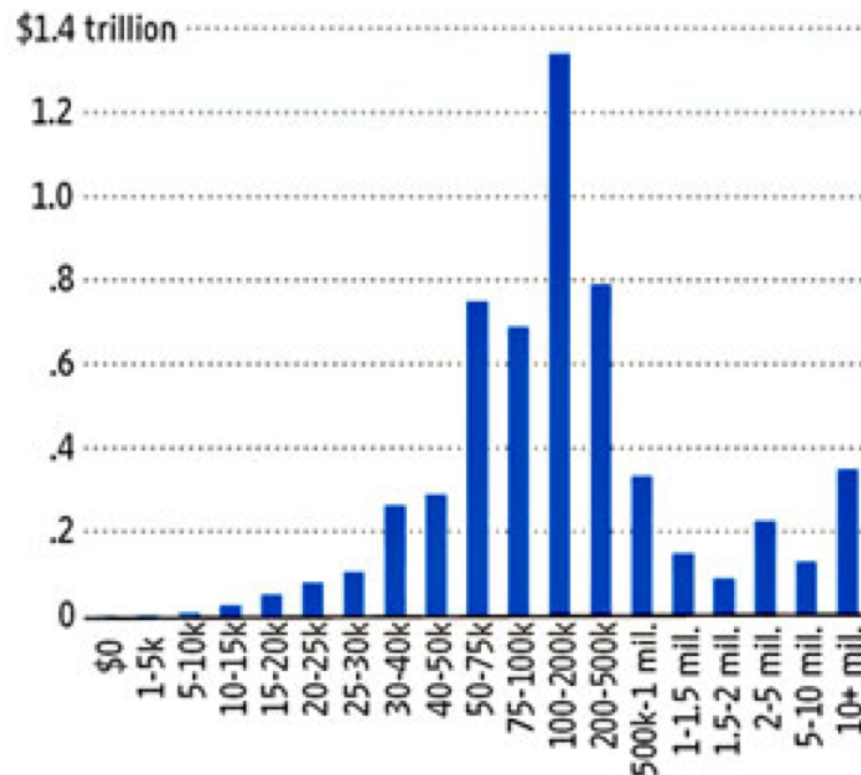|											ME
AK											
								VT	NH	MA	
WA	MT	ND	SD	MN	WI	MI		NY	CT	RI	
OR	ID	WY	NE	IA	IL	IN	OH	PA	NJ		
CA	NV	UT	CO	KS	MO	KY	WV	DC	MD	DE	
	AZ	NM	OK	AR	TN	VA	NC				
		TX	LA	MS	AL	GA	SC				
HI						FL					

Different Aggregations Tell Different Stories

The Middle Class Tax Target

The amount of total taxable income (left scale) for all filers by adjusted gross income level for 2008

Summary: Telling the Truth

- Context
- Size coding – lie factor
- Zero origin
- Inflation adust on dollars
- Per-capita adjust – college degrees (etc.) per 100,000 people
- Size adjust on map (cartogram)
- Aggregations

2. Design Aesthetics (Principles)

- Set of principles to help guide designers

- Maximize data-ink ratio
- Content is king (less is more)
- Careful with 3D
- Avoid "chartjunk"
- Use multi-functioning graphical elements
- Be careful with color
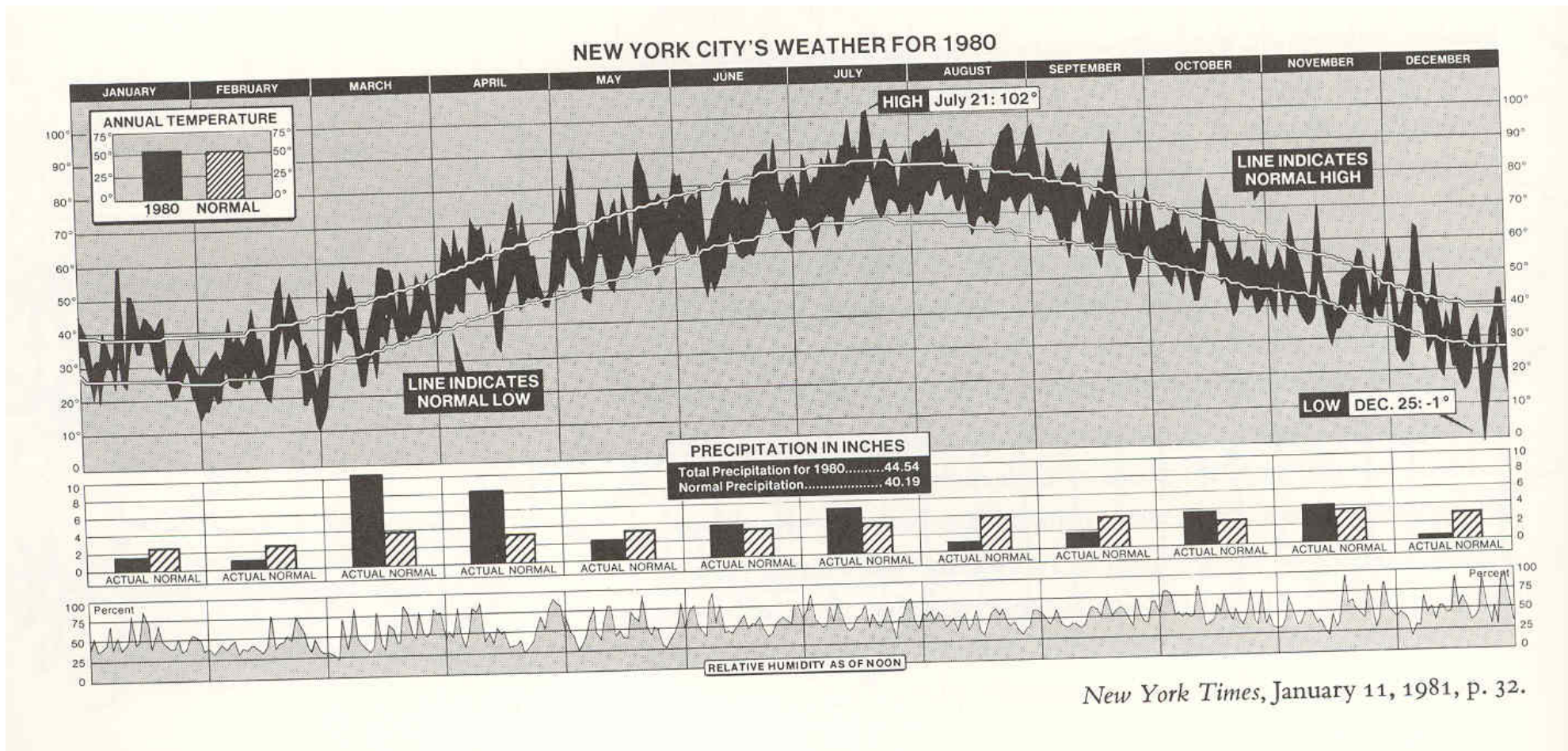
Design Principle: Data-Ink Ratio

- **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

Data – Ink Ratio Example

Good

Bad

Tufte, Vol 1, p. 94

Data – Ink Ratio Example

NEW YORK CITY'S WEATHER FOR 1980

New York Times, January 11, 1981, p. 32.

Outstanding – a lot of information!

Tufte, Vol 1, p. 30

Design Principle: Content is King

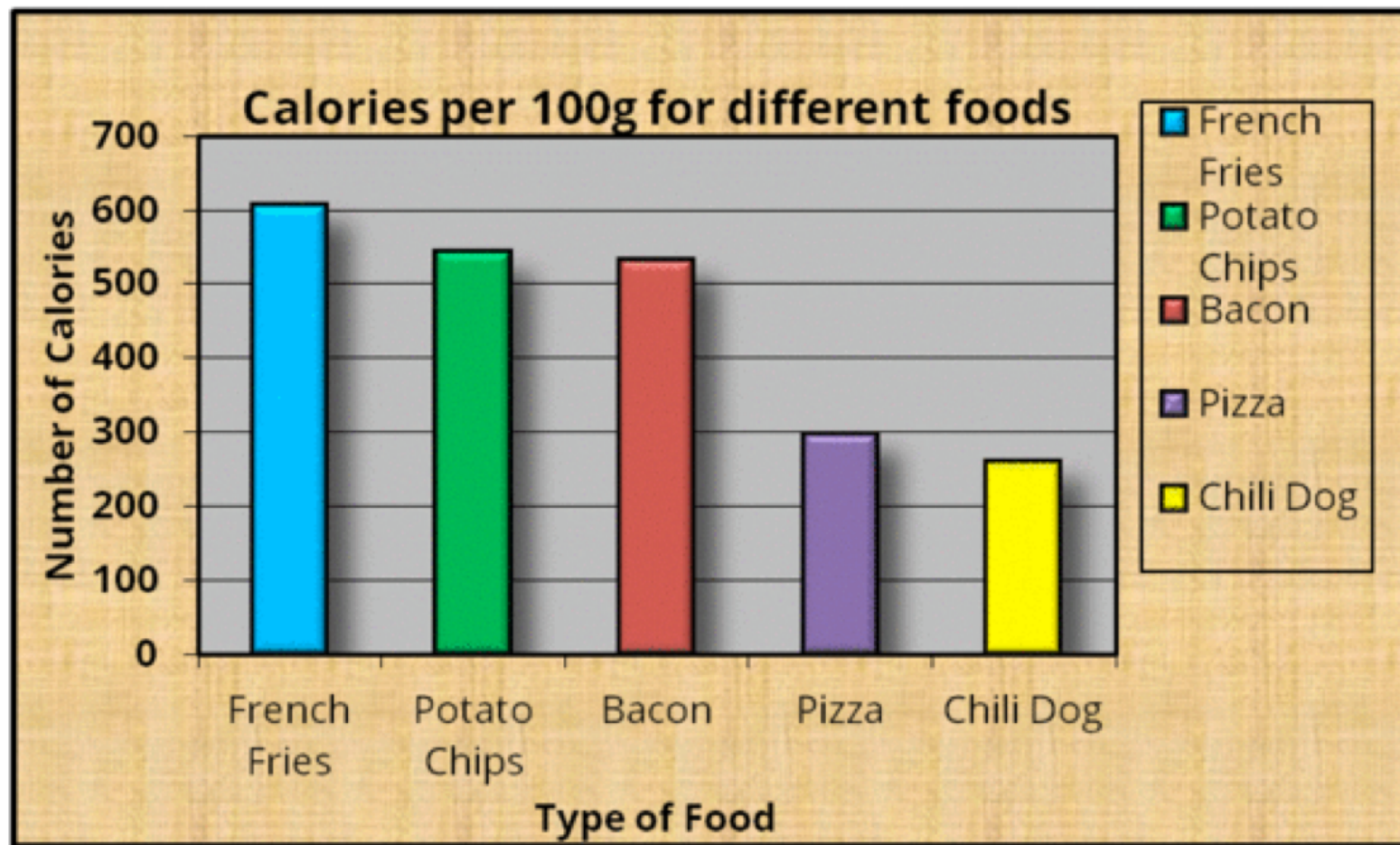- ## Content is king

  - Avoid separate legends and keys -- Just have that information in the graphic

    (**Follow the eyes**……)

  - Make grids, labeling, etc., very faint so that they recede into background

  - Integrate text, chart, graphic, map into a coherent narrative
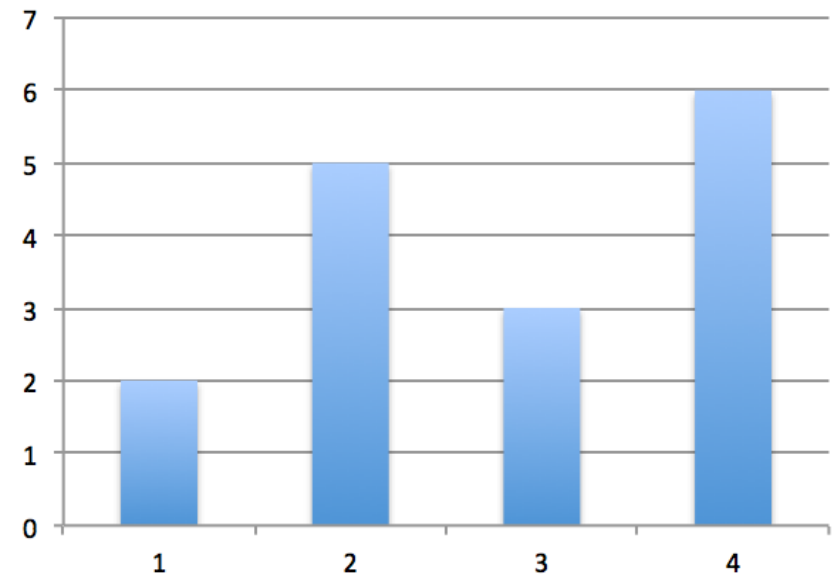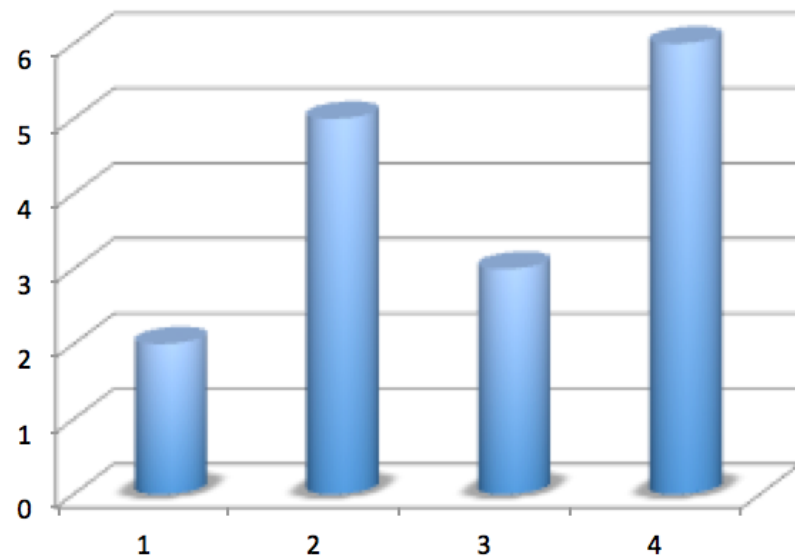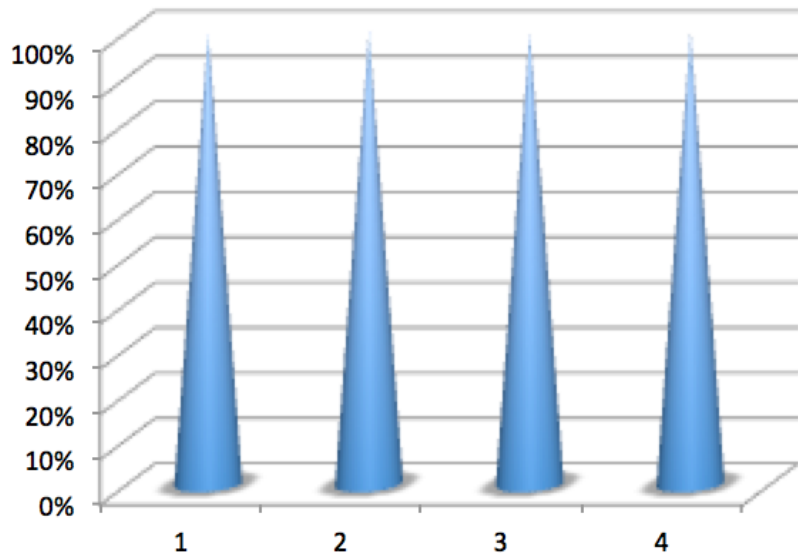
  - Avoid fancy visual effects

Less is More

- http://i.imgur.com/WntrM6p.gif

Calories per 100g for different foods

Content is King: No Fancy Visual Effects

Courtesy Microsoft Excel ☹

Before and After Examples

From the web site of Stephen Few

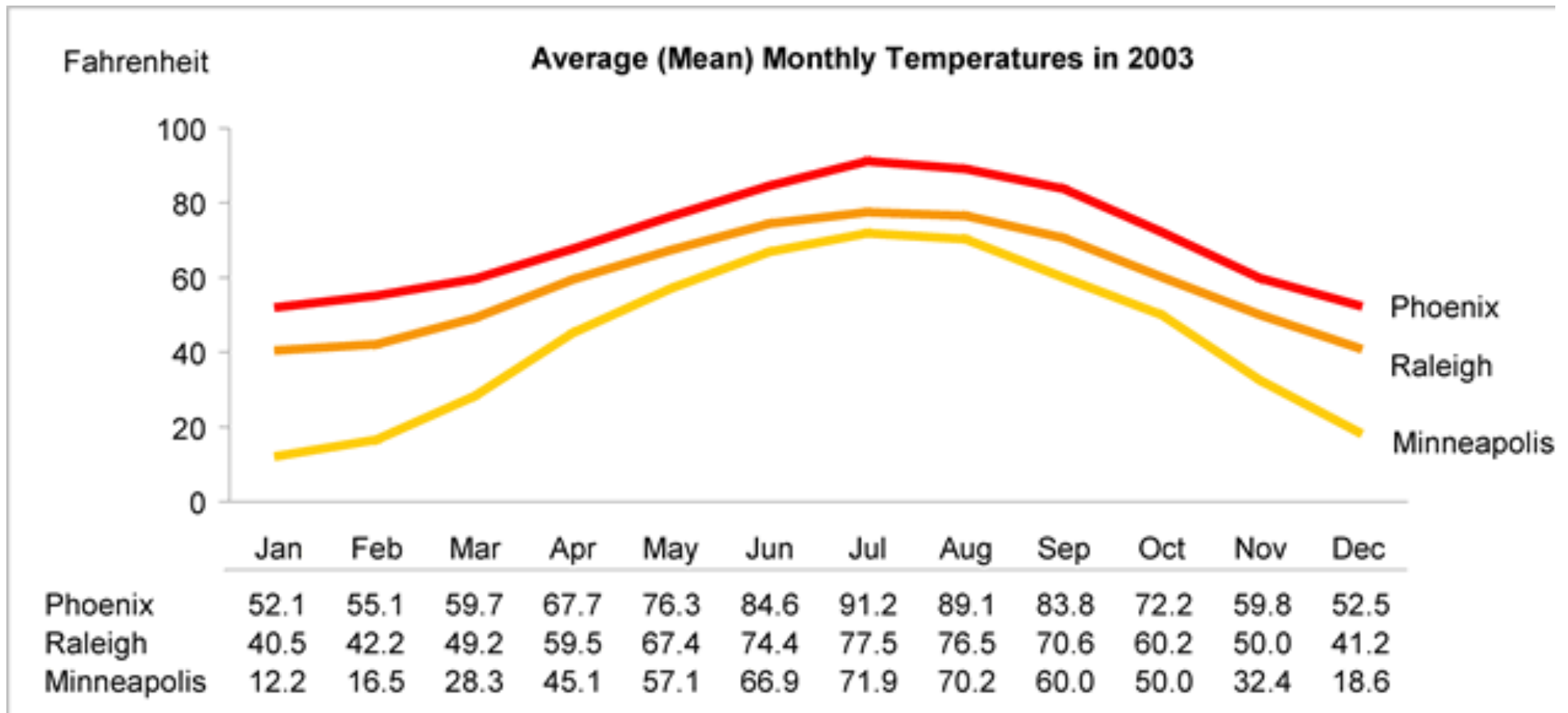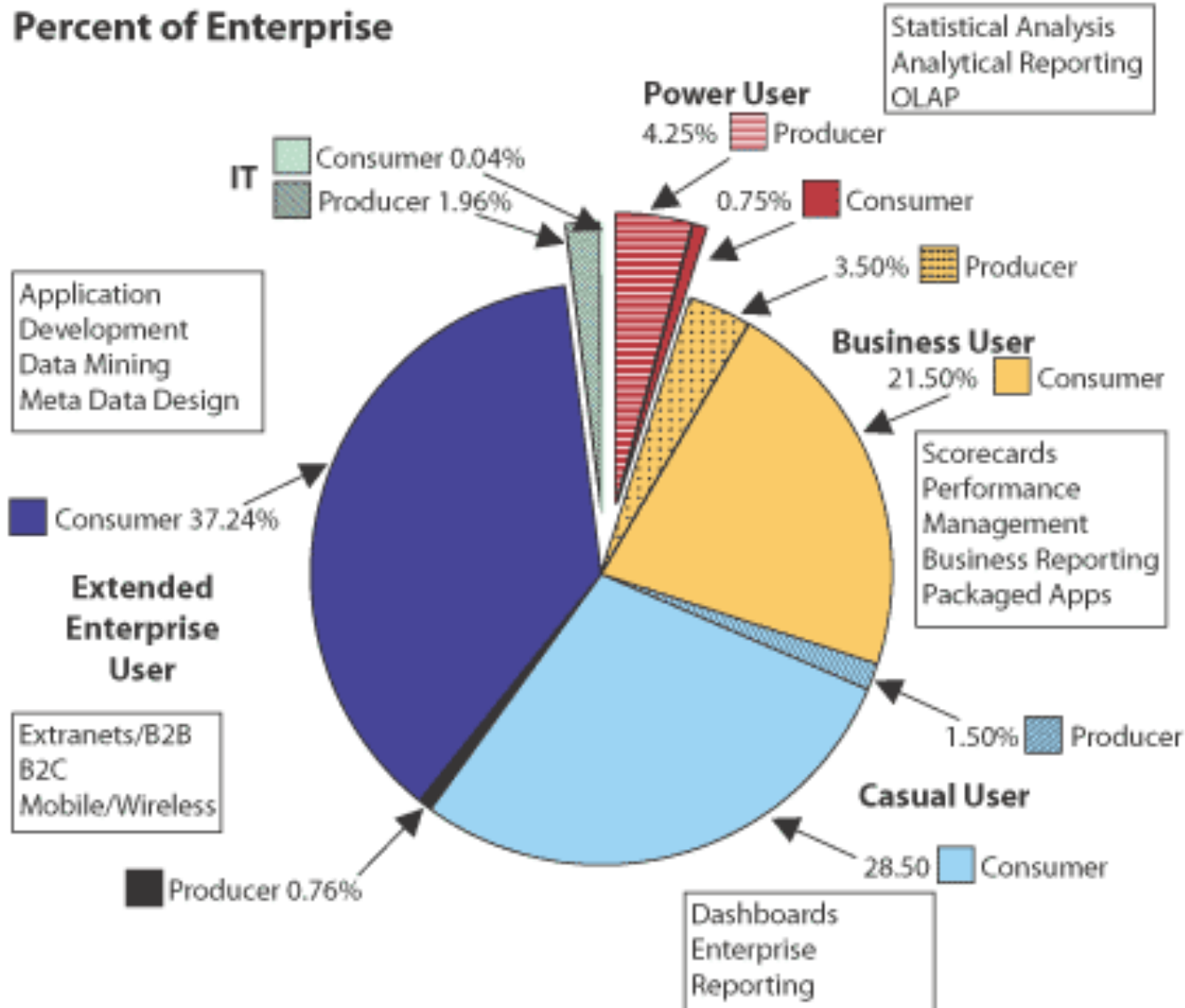http://www.perceptualedge.com/examples.php

Temperatures

How would you redesign?

Average Monthly Temperature

Few's Redesign

Fahrenheit	Average (Mean) Monthly Temperatures in 2003											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Phoenix	52.1	55.1	59.7	67.7	76.3	84.6	91.2	89.1	83.8	72.2	59.8	52.5
Raleigh	40.5	42.2	49.2	59.5	67.4	74.4	77.5	76.5	70.6	60.2	50.0	41.2
Minneapolis	12.2	16.5	28.3	45.1	57.1	66.9	71.9	70.2	60.0	50.0	32.4	18.6
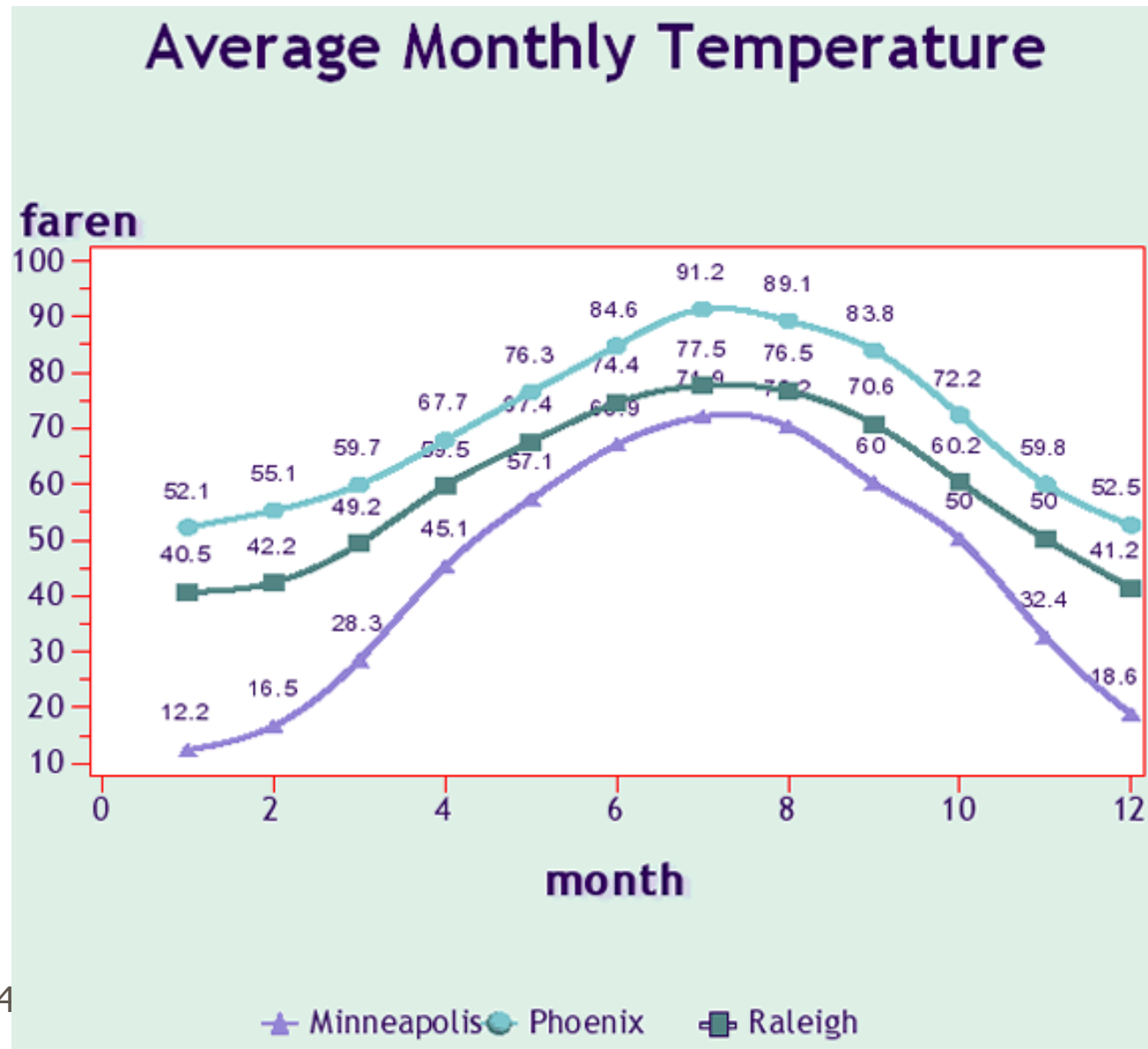
Computer Utilization - Before

Percent of Enterprise

Statistical Analysis
Analytical Reporting
OLAP

Power User
4.25% Producer
0.75% Consumer

IT
Consumer 0.04%
Producer 1.96%

3.50% Producer

Business User
21.50% Consumer

Application
Development
Data Mining
Meta Data Design

Scorecards
Performance
Management
Business Reporting
Packaged Apps

Consumer 37.24%

**Extended
Enterprise
User**

Extranets/B2B
B2C
Mobile/Wireless

1.50% Producer

Casual User

Producer 0.76%

28.50 Consumer

Dashboards
Enterprise
Reporting

After

Percentage of Analytic Computer Usage by Type

Extended Enterprise User (Extranets/B2B/B2C and mobile/ wireless)

Casual User (Dashboards and enterprise reporting)

Business User (Scorecards, performance mgmt, business reporting, and packaged apps)

Power User (Statistical analysis, analytical reporting and OLAP)

IT User (Application development, data mining and meta data design)

■ Consumer   ■ Producer
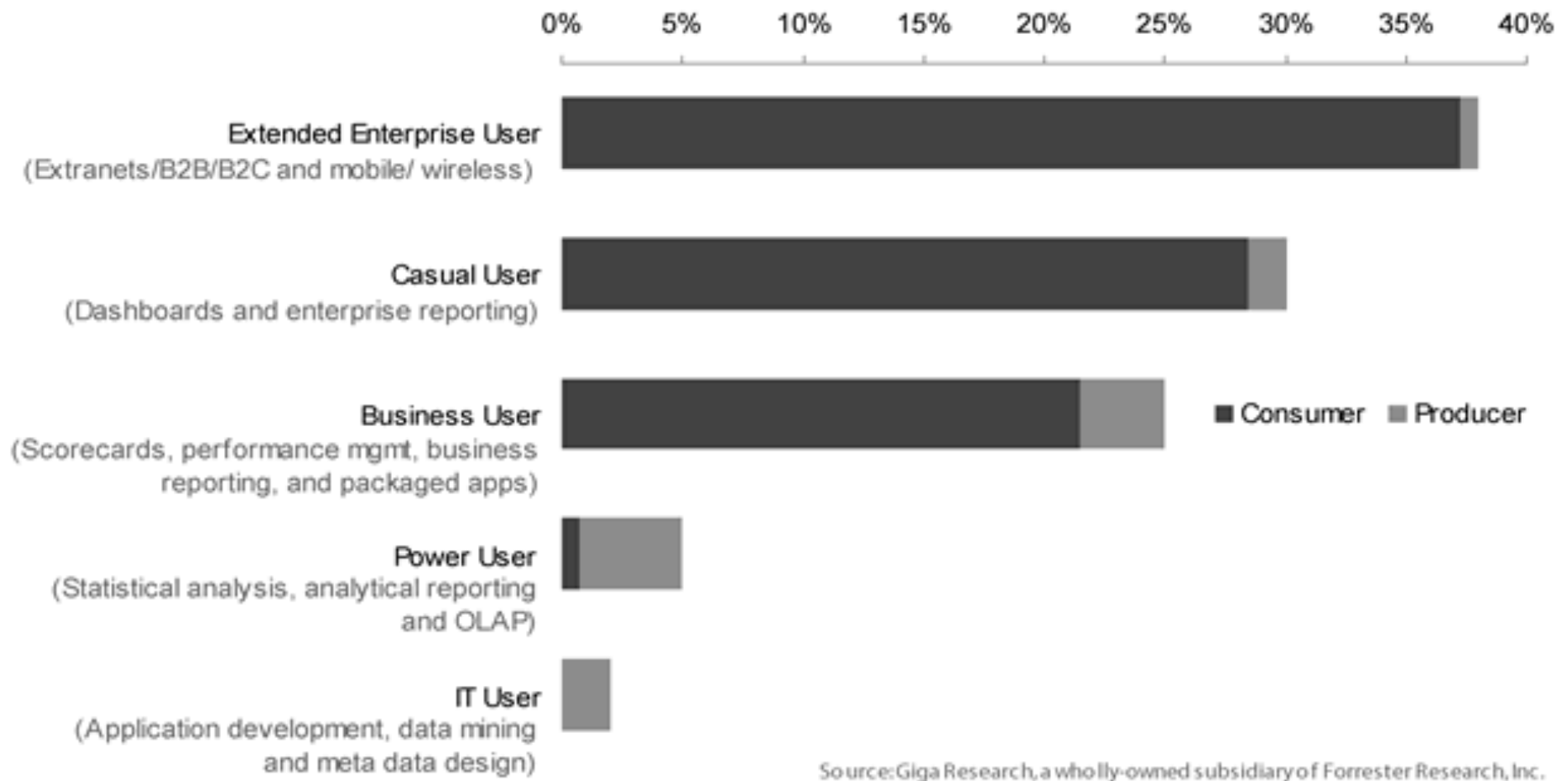
Source: Giga Research, a wholly-owned subsidiary of Forrester Research, Inc.

Attitude Toward USA - Before

Favorable or Unfavorable View of the U.S.

|	
Brazil: % with somewhat or very favorable opinion of the U.S.:	52%
Brazil: % with somewhat or very unfavorable opinion of the U.S.:	32%
Mexico: % with somewhat or very favorable opinion of the U.S.:	64%
Mexico: % with somewhat or very unfavorable opinion of the U.S.:	25%
Britain: % with somewhat or very favorable opinion of the U.S.:	75%
Britain: % with somewhat or very unfavorable opinion of the U.S.:	16%
Germany: % with somewhat or very favorable opinion of the U.S.:	61%
Germany: % with somewhat or very unfavorable opinion of the U.S.:	35%
Russia: % with somewhat or very favorable opinion of the U.S.:	61%
Russia: % with somewhat or very unfavorable opinion of the U.S.:	33%
Poland: % with somewhat or very favorable opinion of the U.S.:	79 %
Poland: % with somewhat or very unfavorable opinion of the U.S.:	11%
South Africa: % with somewhat or very favorable opinion of the U.S.:	65%
South Africa: % with somewhat or very unfavorable opinion of the U.S.:	28%
Kenya: % with somewhat or very favorable opinion of the U.S.:	80%
Kenya: % with somewhat or very unfavorable opinion of the U.S.:	15%
India: % with somewhat or very favorable opinion of the U.S.:	54%
India: % with somewhat or very unfavorable opinion of the U.S.:	27%
Japan: % with somewhat or very favorable opinion of the U.S.:	72%
Japan: % with somewhat or very unfavorable opinion of the U.S.:	26%
South Korea: % with somewhat or very favorable opinion of the U.S.:	53%
South Korea: % with somewhat or very unfavorable opinion of the U.S.:	44%
Egypt: % with somewhat or very favorable opinion of the U.S.:	6%
Egypt: % with somewhat or very unfavorable opinion of the U.S.:	69%
Pakistan: % with somewhat or very favorable opinion of the U.S.:	10%
Pakistan: % with somewhat or very unfavorable opinion of the U.S.:	69%
Turkey: % with somewhat or very favorable opinion of the U.S.:	30%
Turkey: % with somewhat or very unfavorable opinion of the U.S.:	55%
Jordan: % with somewhat or very favorable opinion of the U.S.:	25%
Jordan: % with somewhat or very unfavorable opinion of the U.S.:	75%

After

Current World Opinions About the U.S.A

| Favorable or Somewhat Favorable | Unfavorable or Somewhat Unfavorable | Undecided or Neutral |
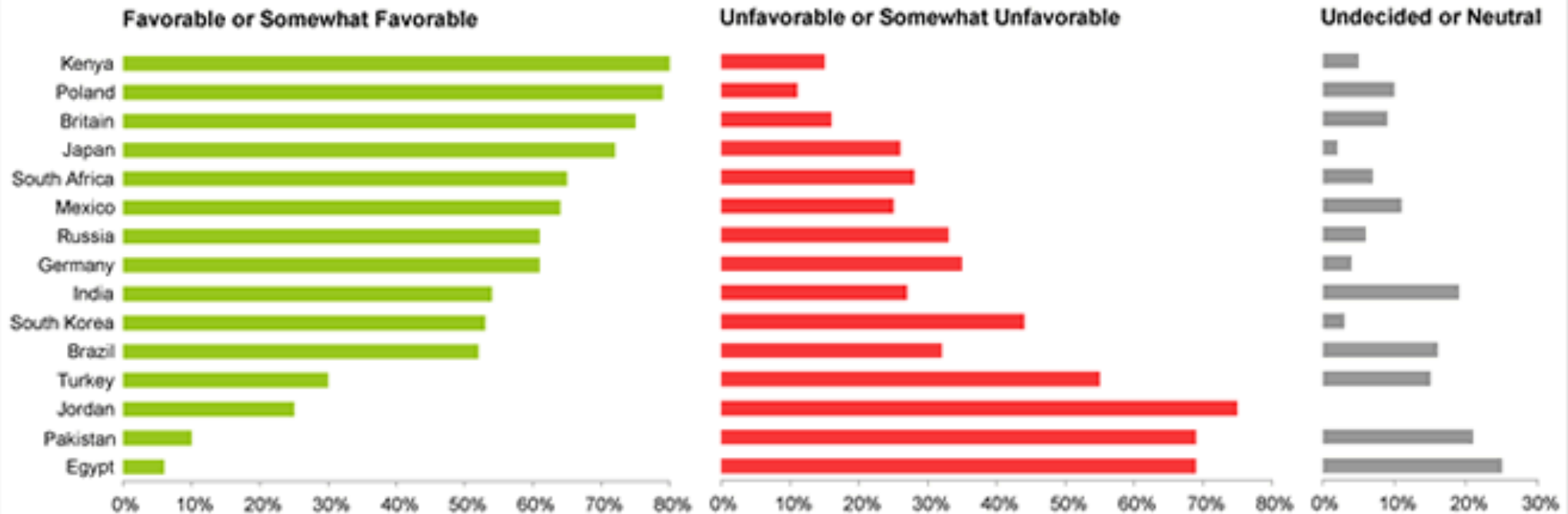
Source: 2004 study conducted by the Pew Center.

Market Share – How would you redesign?
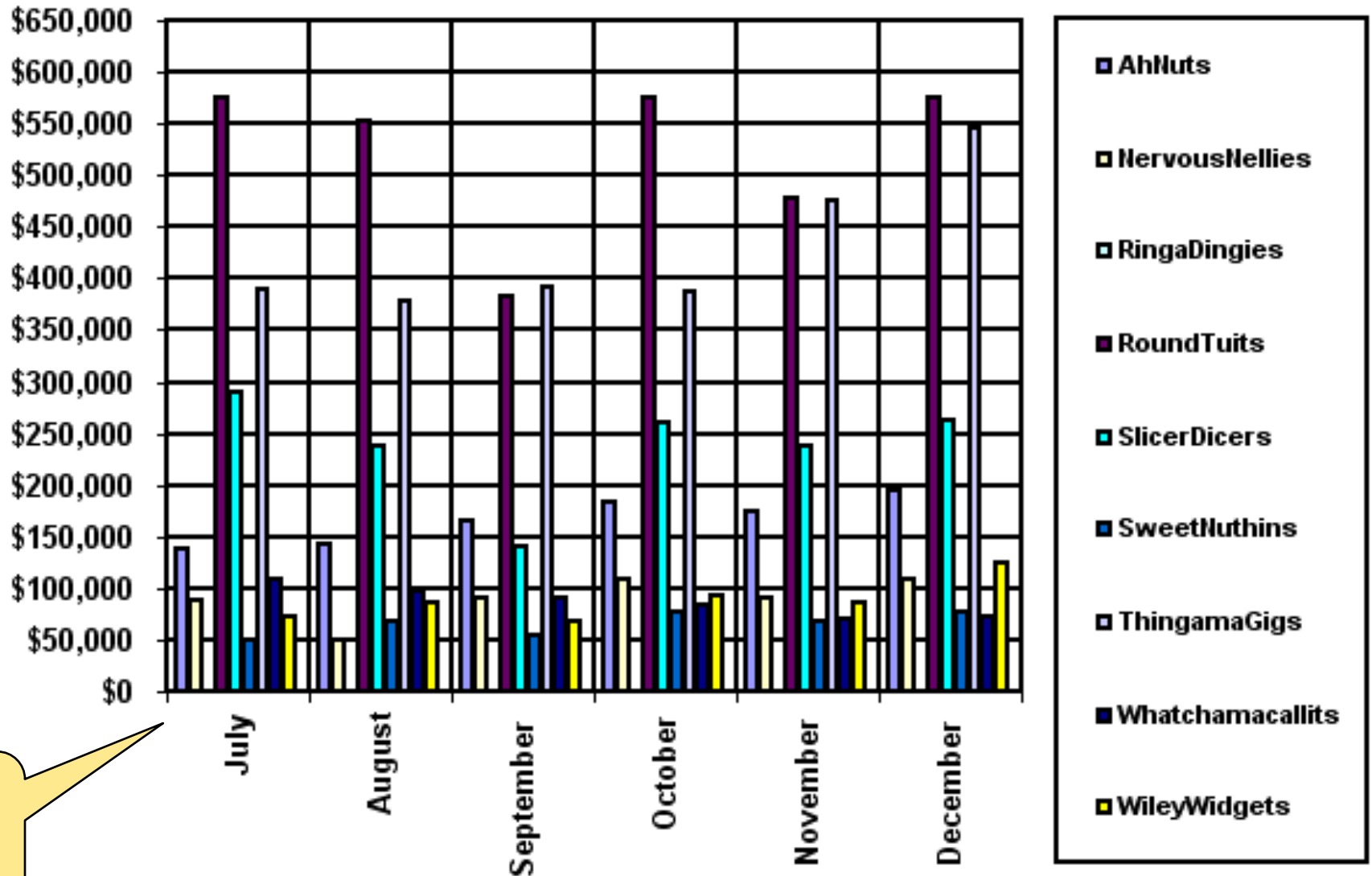
Few's Redesign

Small multiples
Clean graphics

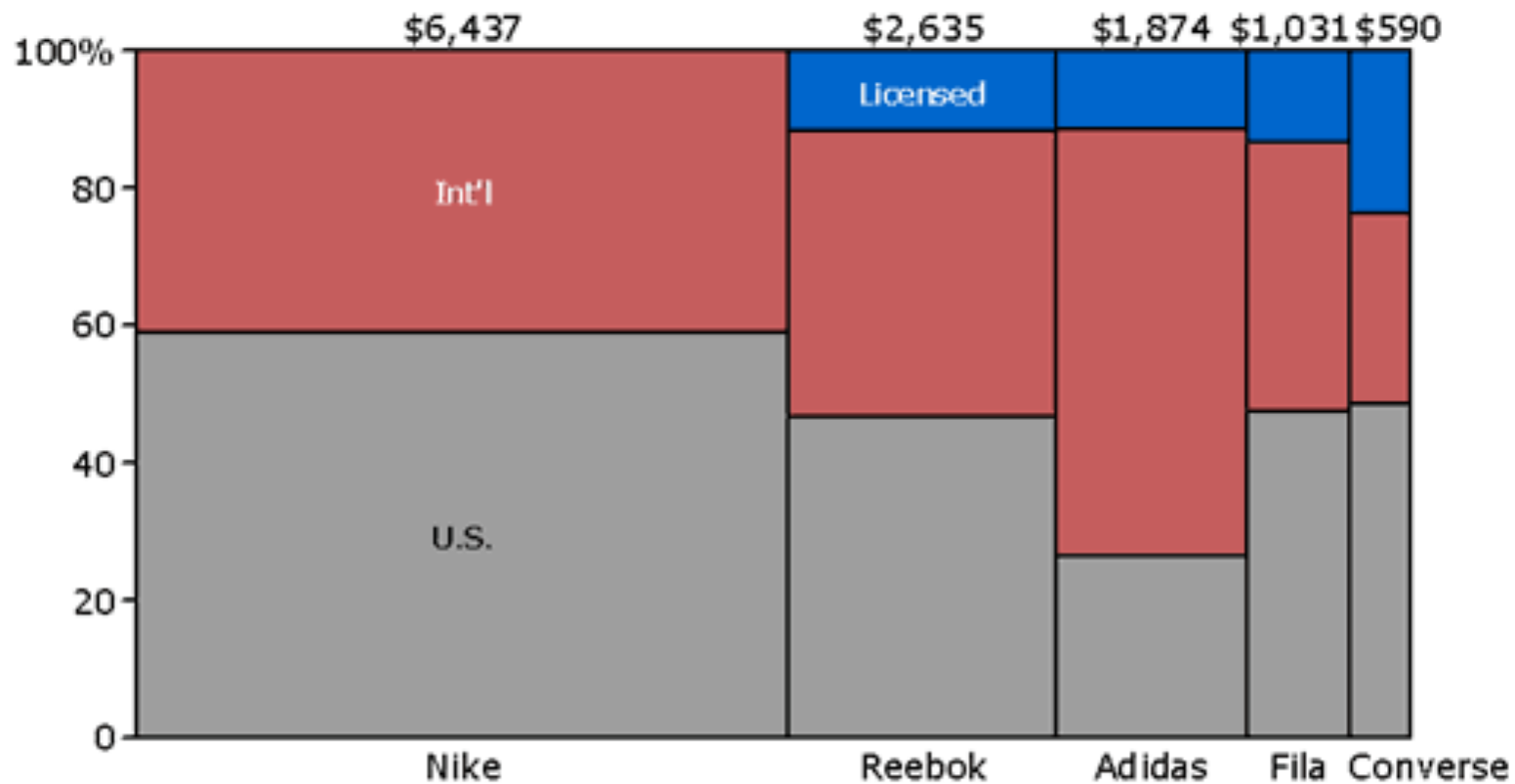CS 4460

Survey Results – What's Wrong?

After

Employee Responses Regarding Company Goals
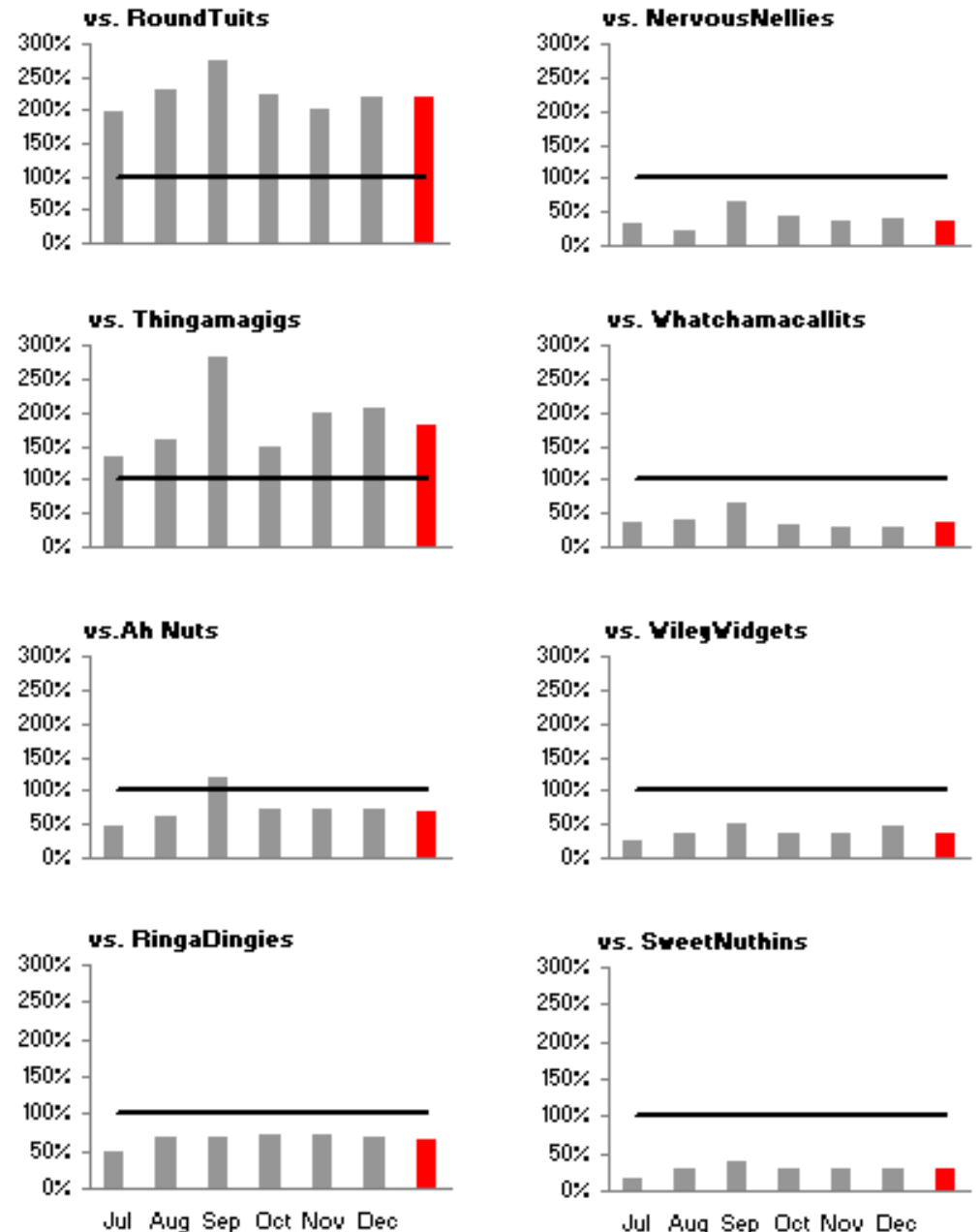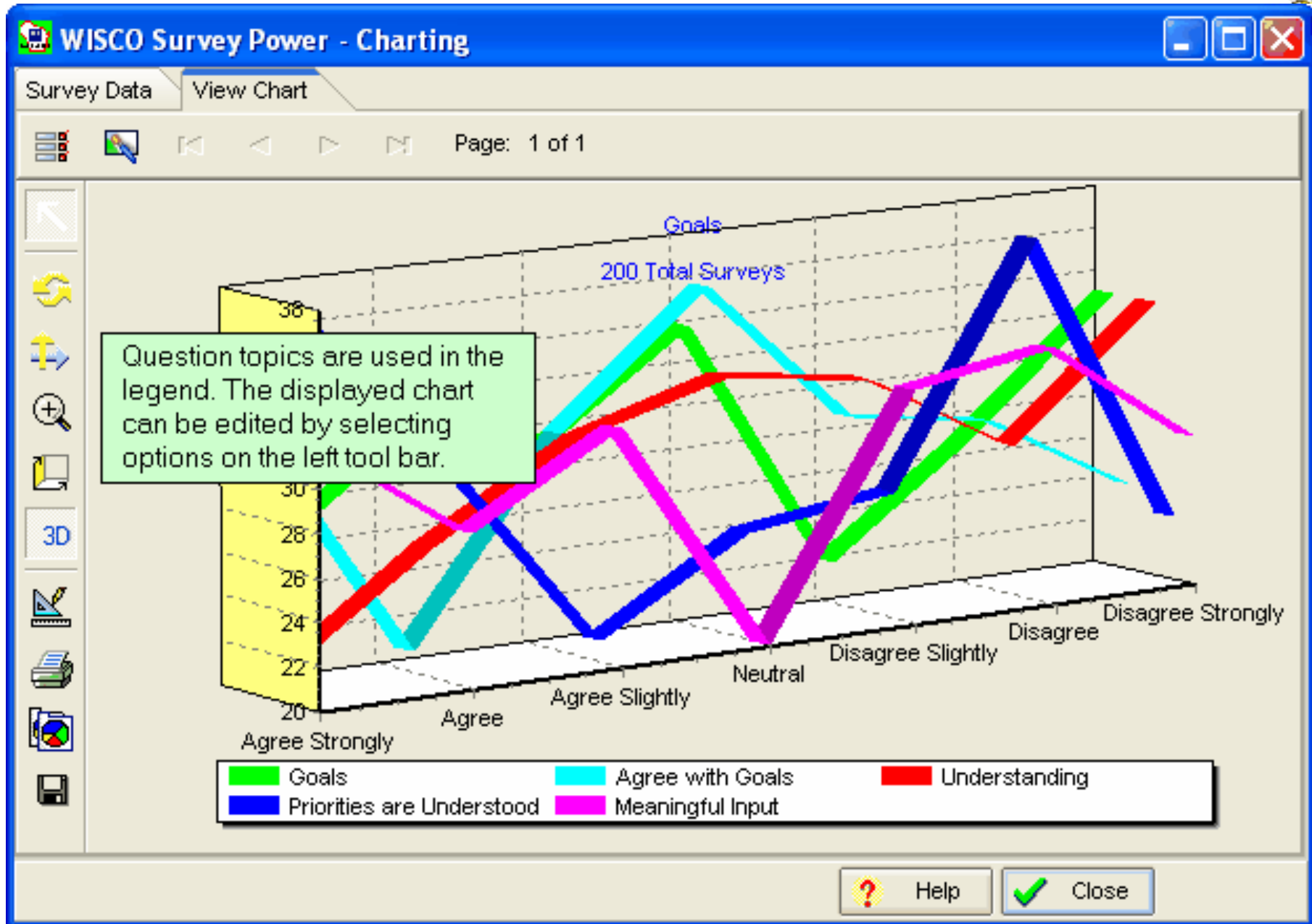
Market Share by Country - Before

Nike dominates its top four competitors with a mix of U.S. and international sales.

Footwear Sales in $ Million

After- Easier to Answer Questions

Total Footwear Sales by Company

Percent of Total Sales / Millions

Companies: Nike, Reebok, Adidas, Fila, Converse

Bottom chart categories: US Sales, International Sales, Licensed Sales

Total Footwear Sales by Channel

Percent of Total Sales / Millions

Channels: U.S. Sales, International Sales, Licensed Sales

Bottom chart legend: Nike, Reebok, Adidas, Fila, Converse
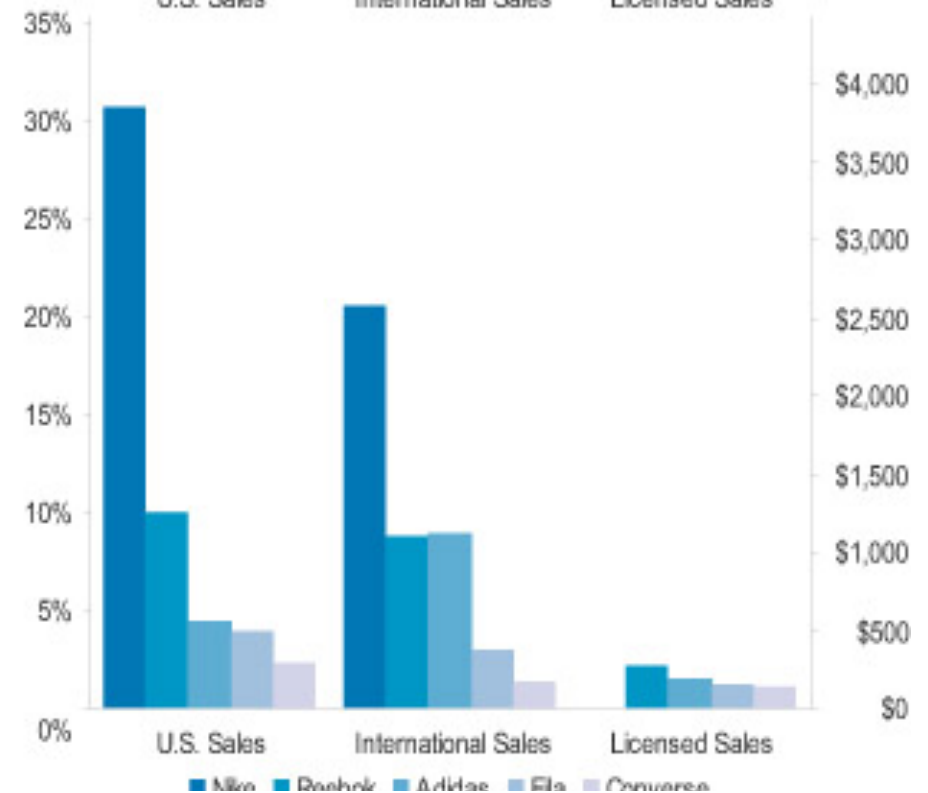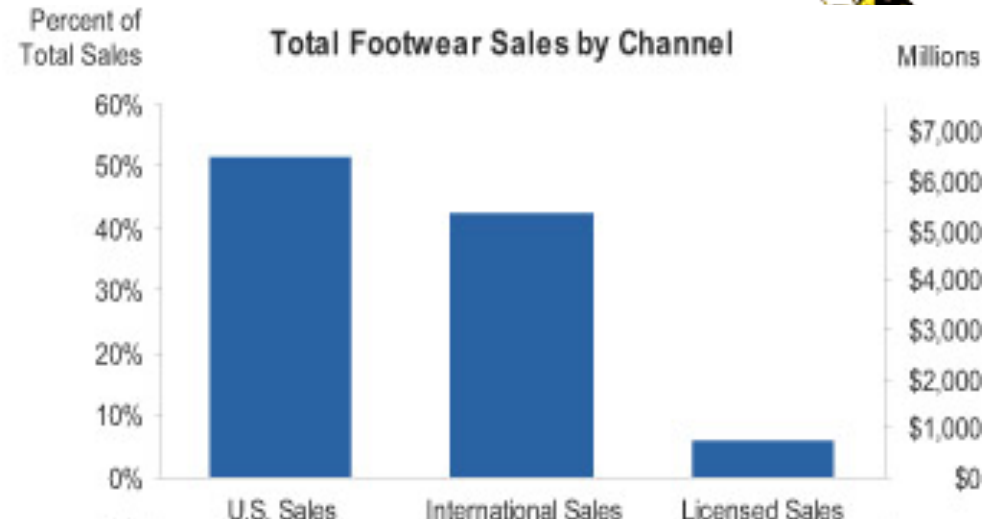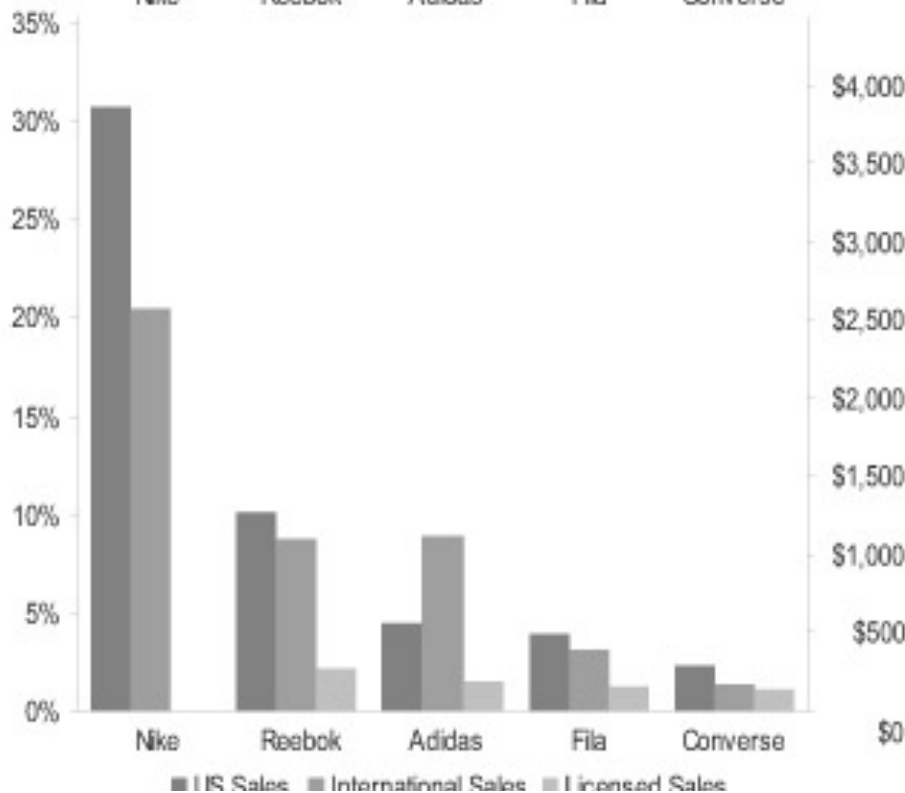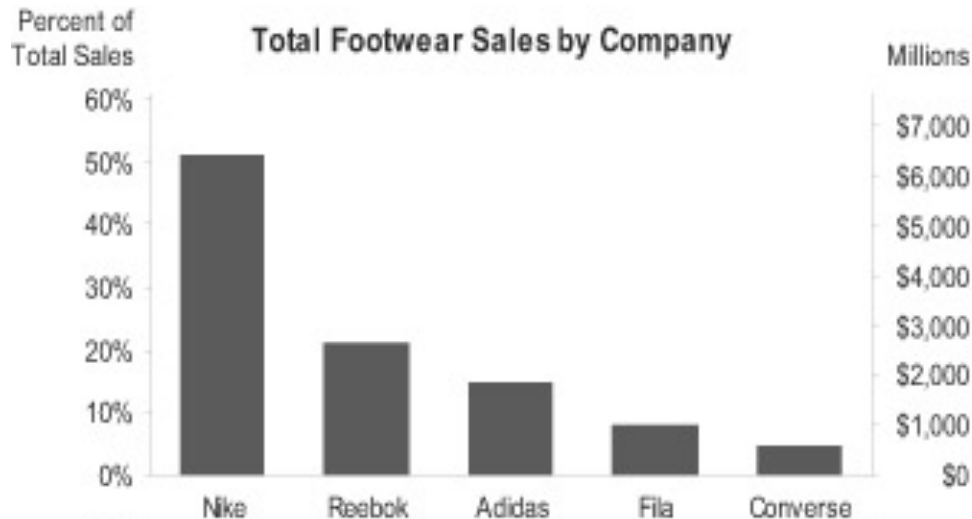
End of Before and After Examples

Design Principle: Careful with 3D

- Be careful with 3D (JDF)
- BTW how many variables?

Example from
http://www.perceptualedge.com/blog/?p=978
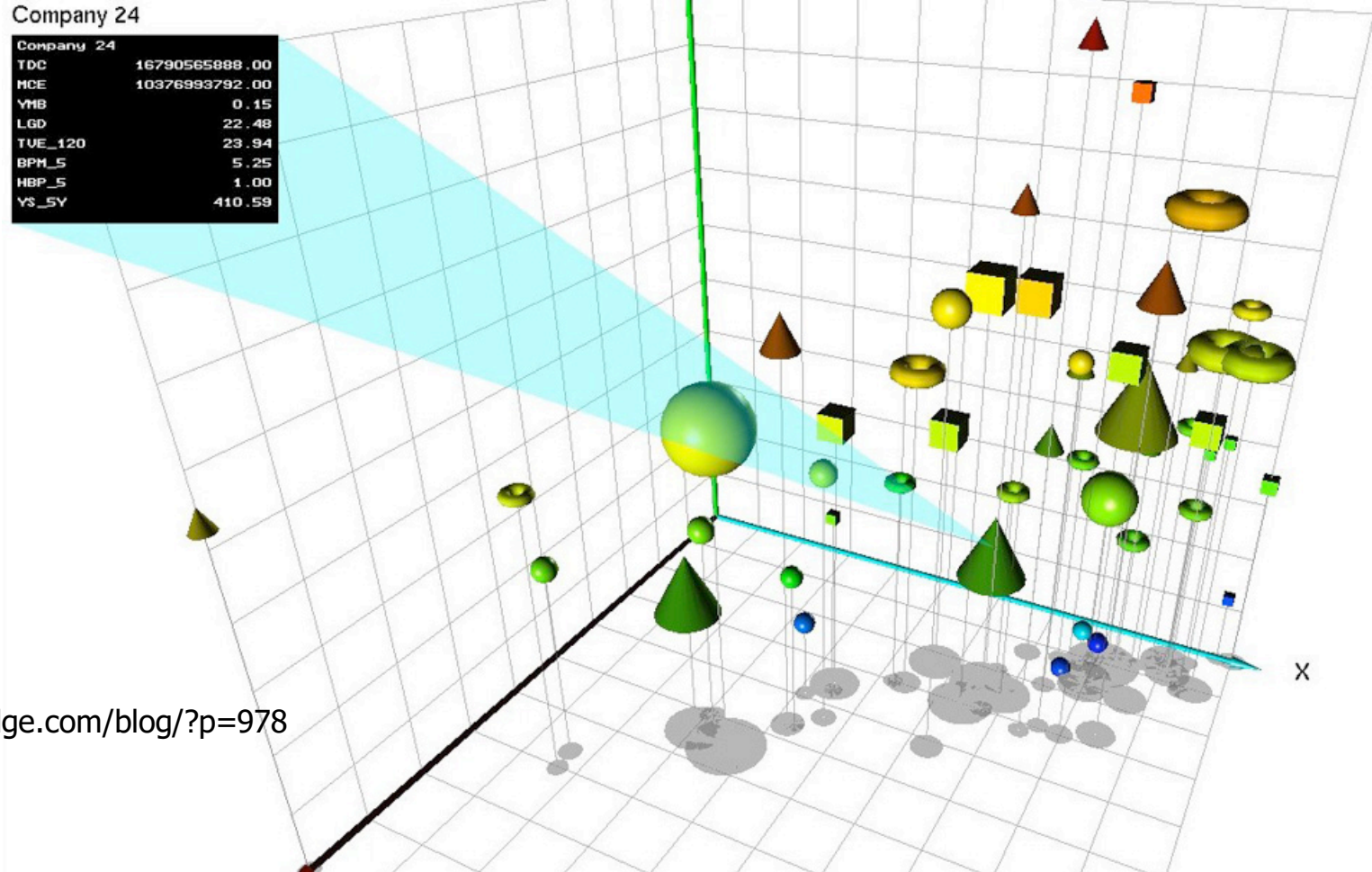
Survey Results – What's Wrong?
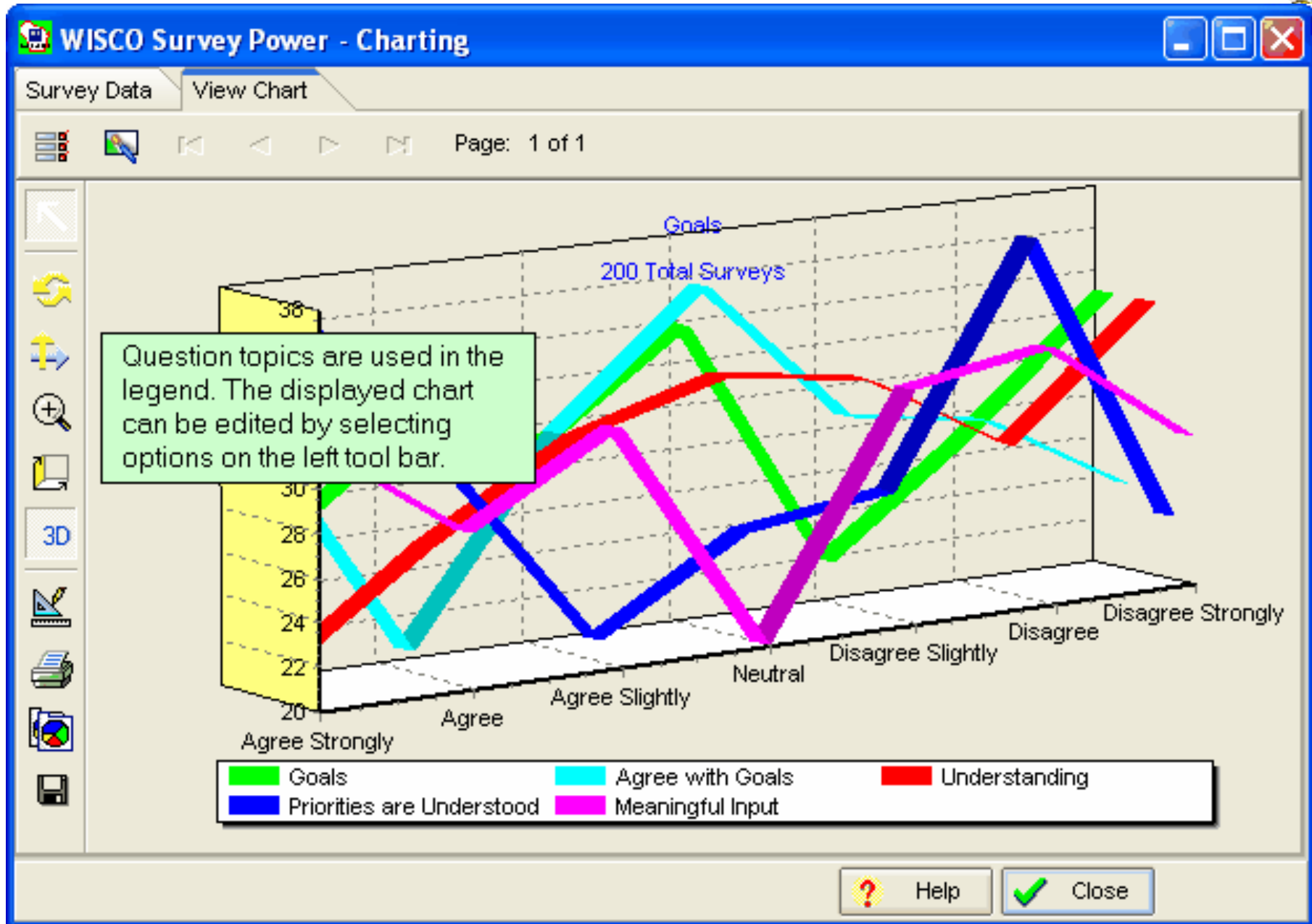
Design Principle: No Chartjunk

- **Avoid chartjunk**
  - Extraneous visual elements that detract from message

Chart Junk

A classic

DIAMONDS *WERE* A GIRL'S BEST FRIEND
Average price of a one-carat D-flawless

$60,000
$50,000
$40,000
$30,000
$20,000

1978 1979 1980 1981 1982

Tufte, Vol 2, p.34

Chart Junk

Chart Junk

Chart Junk!!

Markets — **Chart of the Day**

The price is not right

(house-price-to-rent ratio, 2013:Q2, deviation from historical average)

Undervalued (−)   Overvalued (+)

Japan, Greece, Germany, Portugal, Slovenia, Czech Republic, Slovakia, Iceland, Switzerland, United States, Italy, Ireland, Korea, Luxembourg, Netherlands, Israel, Austria, Denmark, Spain, Sweden, United Kingdom, France, Finland, Australia, Belgium, Norway, New Zealand, Canada

BUSINESS INSIDER

http://www.businessinsider.com/chart-over-and-undervalued-markets-2013-11

And Still More Chart Junk
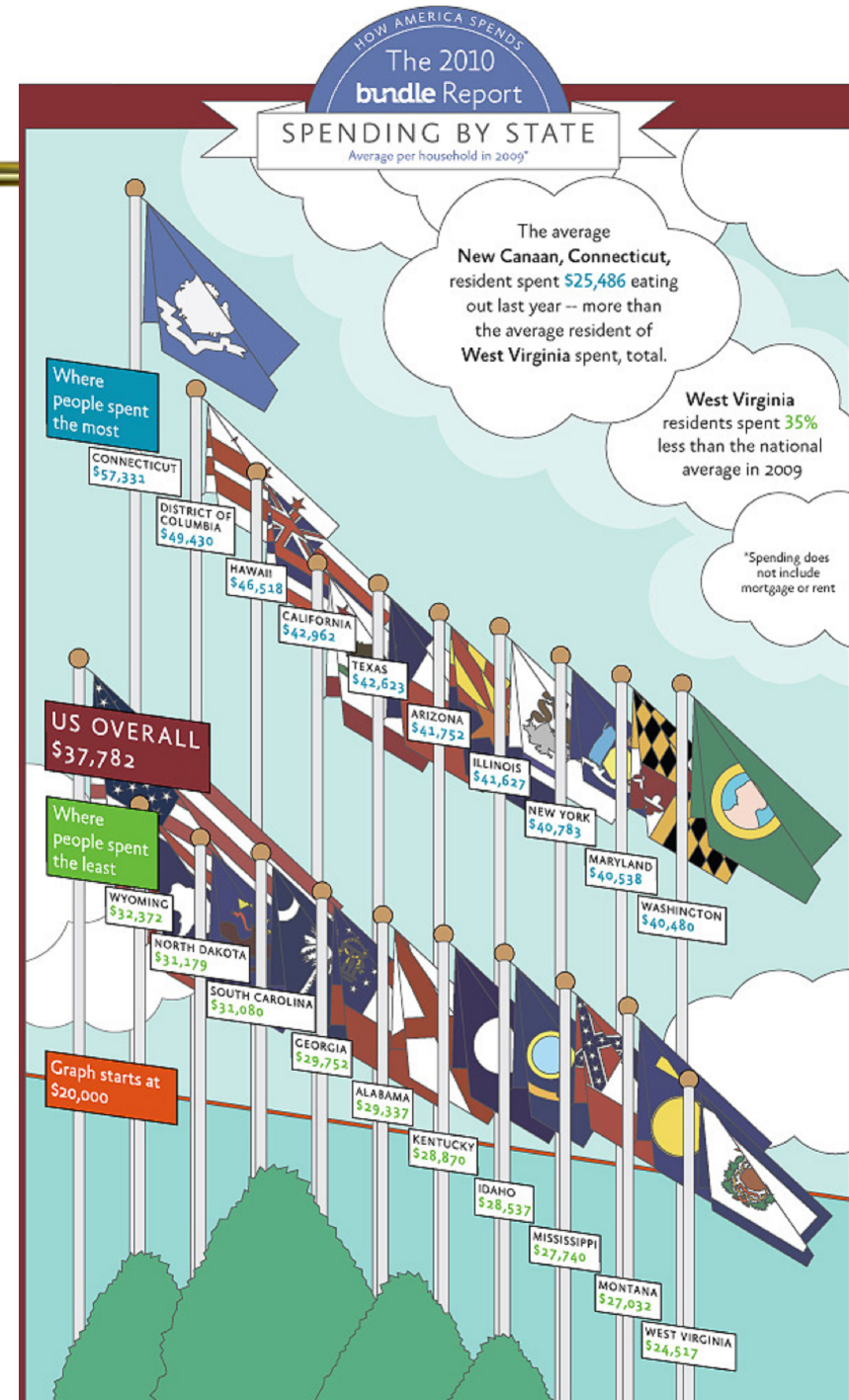
Junk Charts Blog

Great place
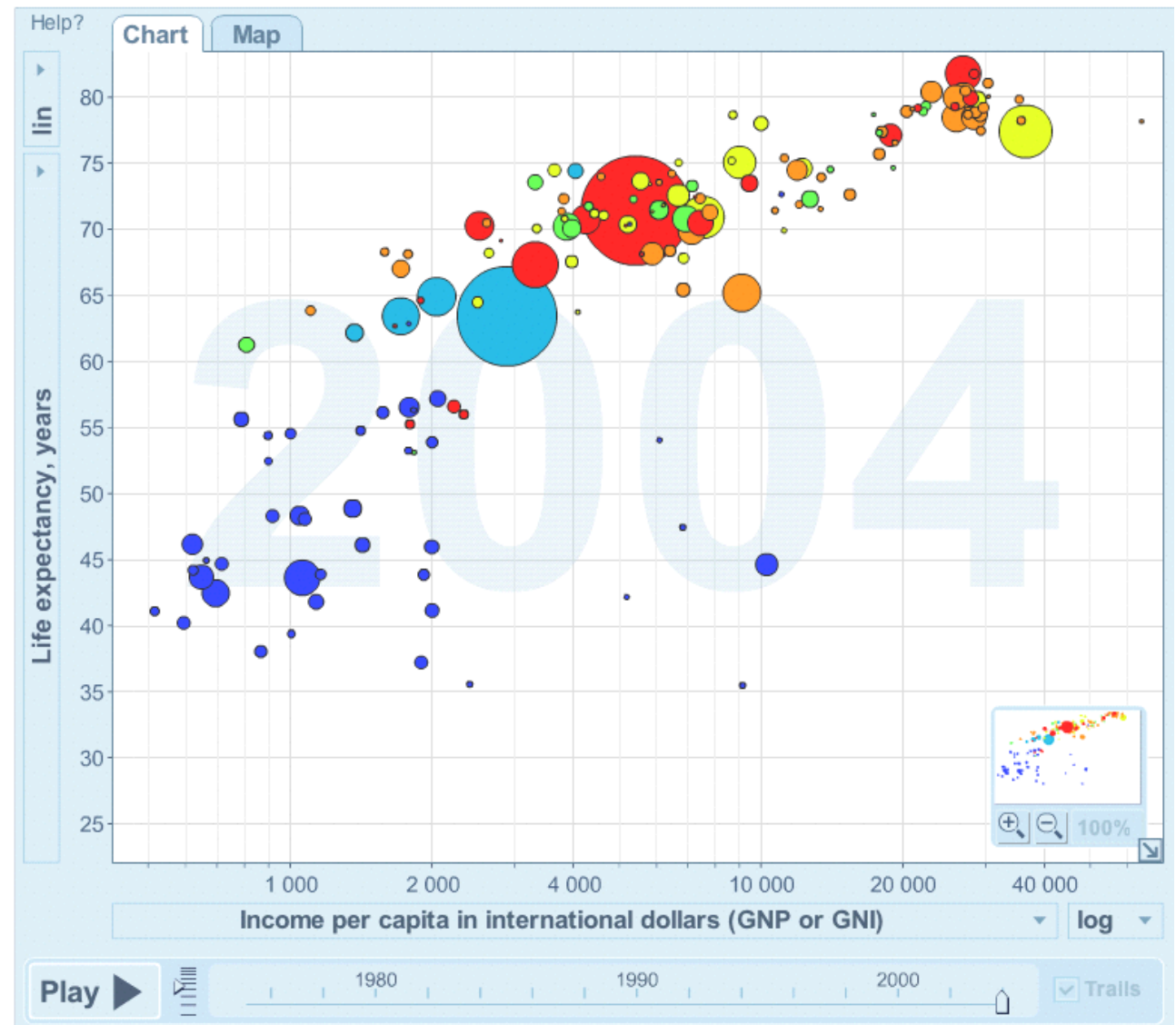to browse!

Design Principle: Multifunction

- **Use multi-functioning graphical elements (macro/micro readings)**
  - Graphical elements that convey more than one piece of information
  - They serve double-duty

Gapminder: Multi-functioning Elements

Bubbles encode:
 Income
 Life expectancy
 Population
 Continent

Multi-functioning graphical elements

0 | 9 = 900 feet

0	98766562
1	97719630
2	69987766544422211
3	876655412099551426
4	99988443319294333611
5	97666666554422210097731
6	898665441077761065
7	98855431100652108073
8	653322122937
9	3776554421000493
10	0984433165212
11	4963201631
12	45421164
13	47830
14	00
15	676
16	52
17	92
18	5
19	39730

One volcano is 200 feet high

BTW what could be improved here?

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

BTW Multi-functioning just means conveying more than one piece of information

One volcano is 13,800 feet high

19 | 3 = 19,300 feet

Tufte, Vol 1, p 140

Design Principle: Careful with Color!

- To differentiate information marks (red states, blue states)

- To show continuous variation of a value (as in temperature range)

- To call attention/highlight
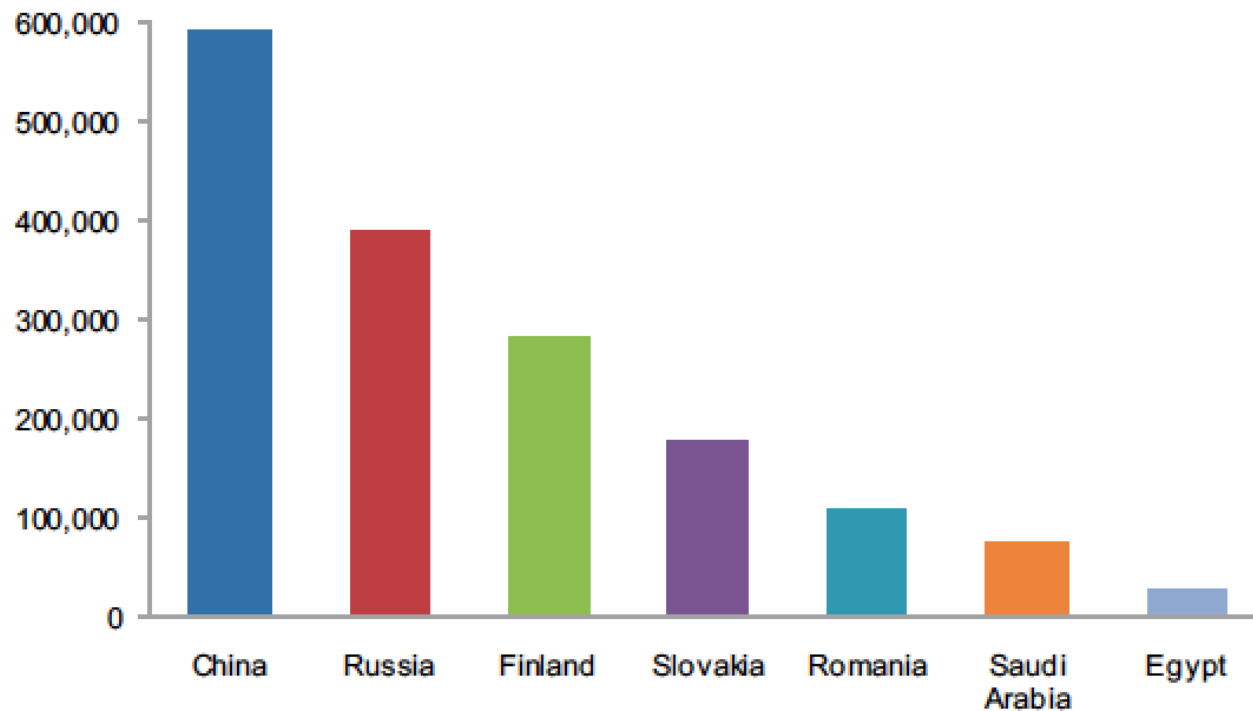
Stephen Few's Color Rules

- Use color ONLY to help communicate
- Worry about color surround
- Color contrast ensures visibility of info
- Use different colors ONLY for different meanings
- Use soft colors for most info, bright and/or dark colors to call attention
- To encode quantitative variable, use single or closely related hues
- Use very light colors (or gray) for non-data info
- Remember the color blind
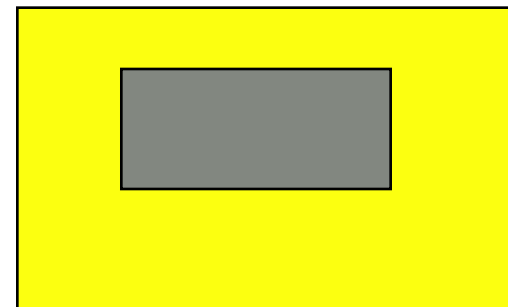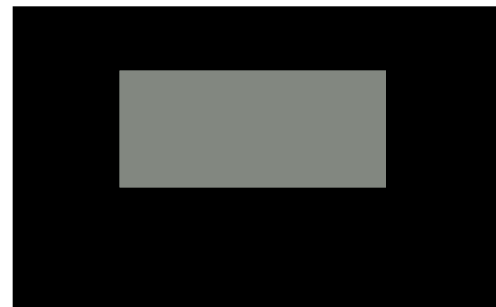- Do NOT use fancy visual effects
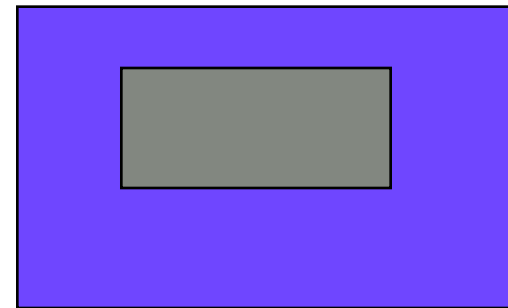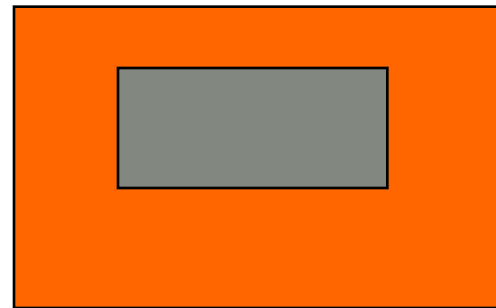
Color ONLY to communicate

Of what use is color in this bar chart?

CSE 4470

Color Surround
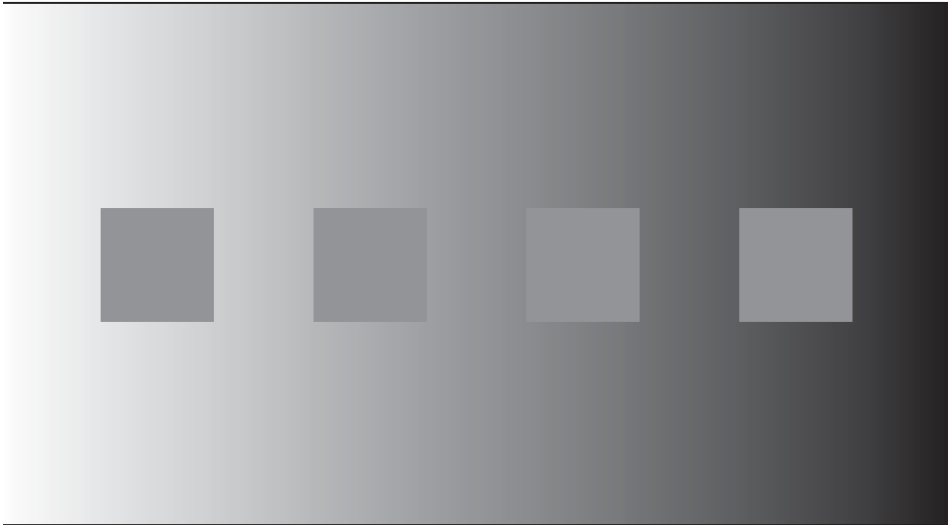
Color Contrast for Visibility

Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?
Hello, here is some text. Can you read what it says?

Color Contrast for Visibility

	Laptops	Desktop PCs	Harddisks	Flash Memory	Screens	Keyboards	Printers	Scanners	PDAs	Projectors	Cameras
California	-2,216	4,497	884	3,252	8,564	3,418	6,582	-3,891	2,333	1,356	5,450
Colorado	3,410	0	2,338	2,676	1,567	367	1,361	3,249	828	1,272	-141
Connecticut	0	0	4	0	2,998	1,219	673	0	0	610	989
Florida	0	0	583	765	2,305	940	1,737	2,727	0	1,344	497
Illinois	0	0	9,384	3,405	4,362	2,331	4,495	3,464	1,366	0	1,569
Iowa	377	0	412	6,577	210	4,487	872	189	5,622	0	3,466
Louisiana	0	2	1,455	1,501	1,138	0	920	1,349	0	0	991
Massachusetts	0	0	-297	0	12,489	707	0	0	0	492	367
Missouri	0	0	1,224	-169	986	367	1,011	1,125	-210	-45	-685
Nevada	0	875	374	884	-45	11,934	410	210	4,497	-10,978	4,355
New Hampshire	1,009	0	-166	0	897	376	0	0	0	882	-43
New Mexico	0	-208	-161	377	902	0	369	-683	0	0	202
New York	0	0	-17,034	0	8,564	3,254	0	0	3,405	2,326	7,589
Ohio	820	0	2,671	495	827	2,305	355	-134	2,724	0	705
Oklahoma	0	2,722	492	495	1,274	0	2,305	700	-15,034	0	571
Oregon	-139	357	818	708	395	487	2,671	834	2,310	2,727	1,269
Texas	0	1,356	3,415	808	5,452	0	2,336	1,567	0	0	829
Utah	1,130	1,006	1,501	1,009	678	-156	1,222	979	367	-208	904
Washington	0	768	488	2,998	1,661	1,452	1,418	564	918	-5	1,141
Wisconsin	497	0	1,418	1,450	576	916	767	1,740	2	0	1,337

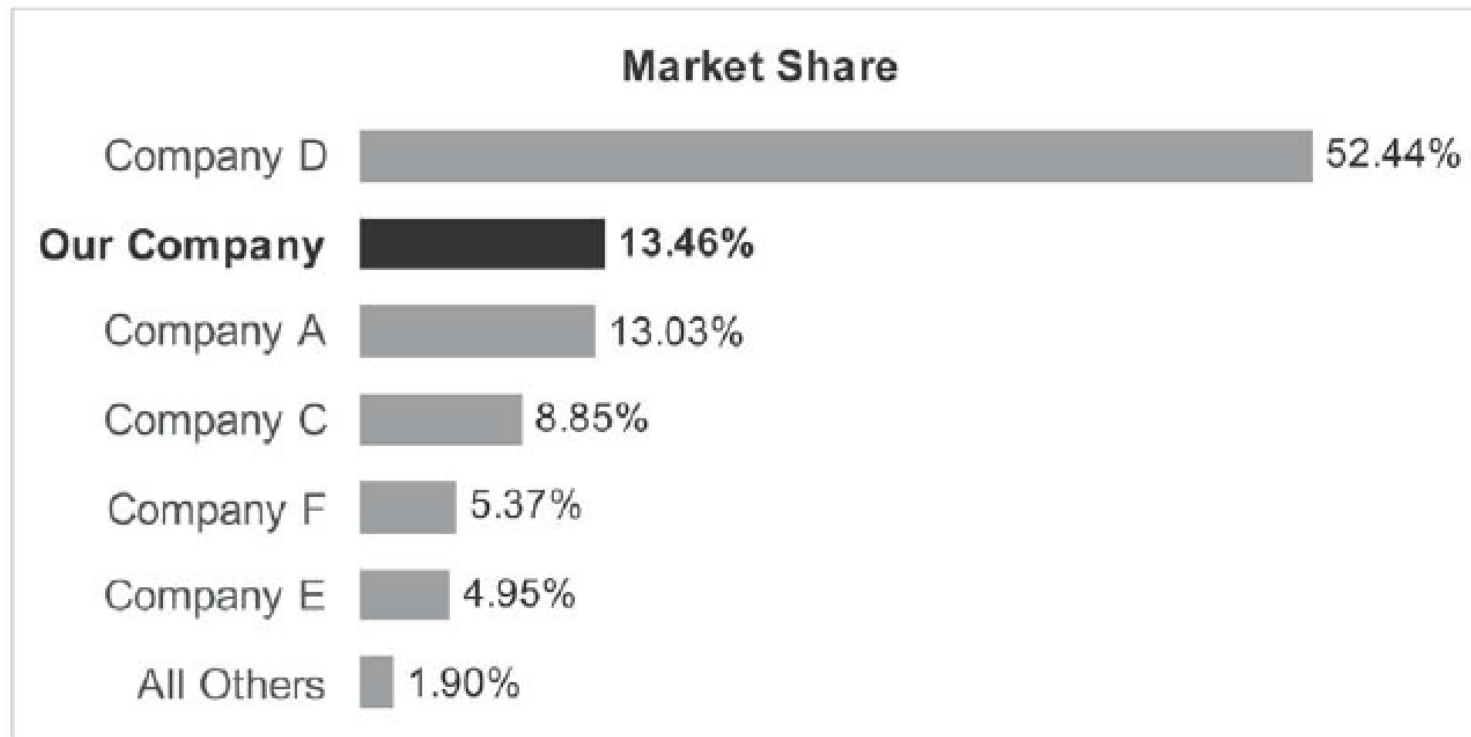# Different Colors ONLY for Diff. Meanings

Which is
better?
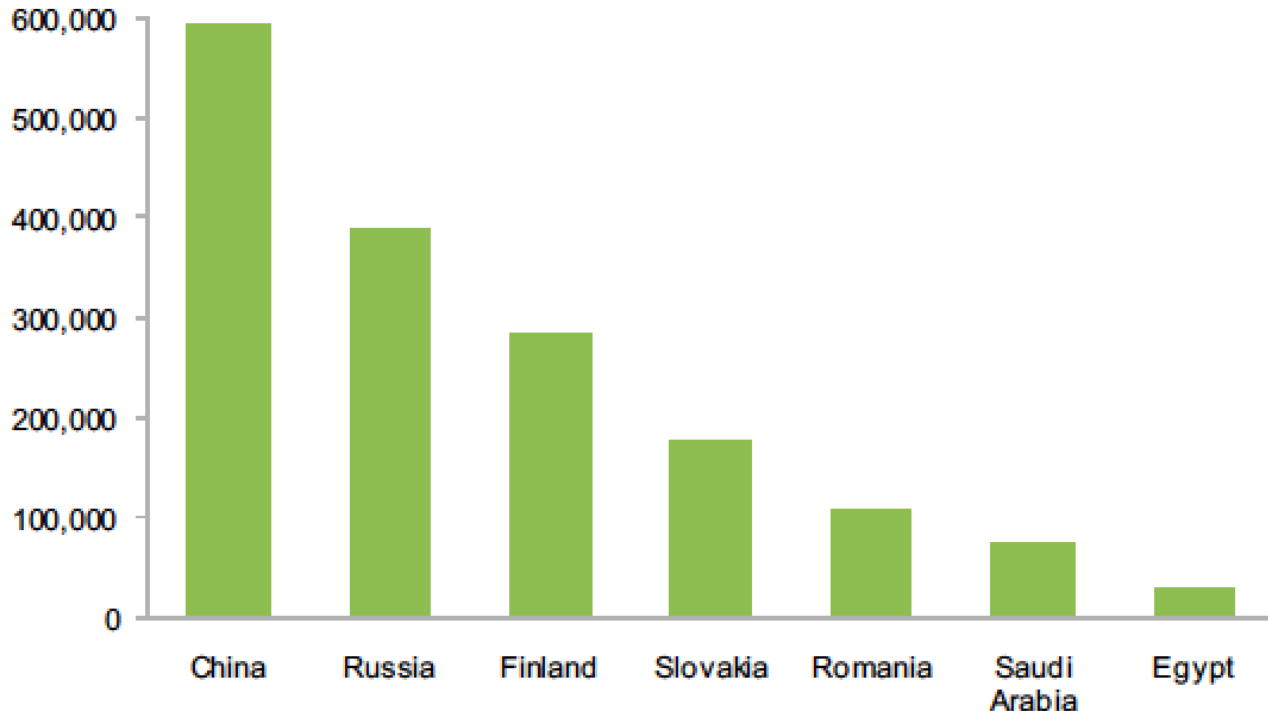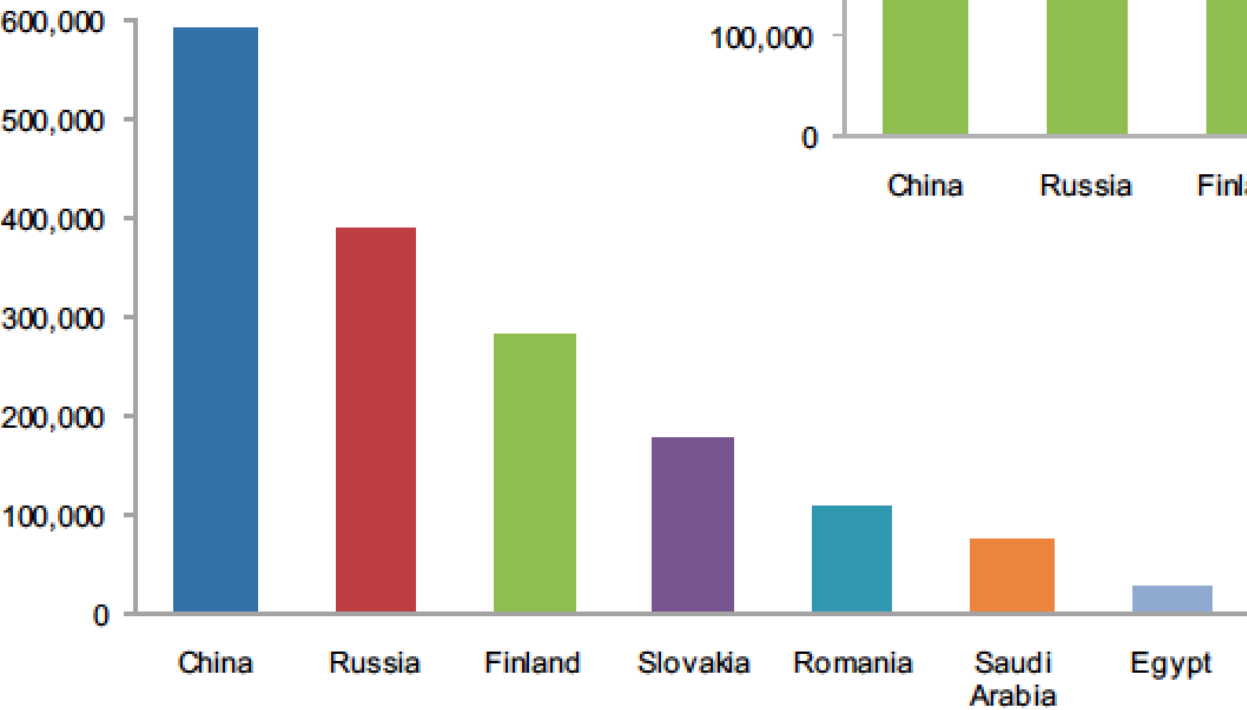
Anything
better still?

Mostly Soft Colors

- Use mostly soft colors, with strong contrasting colors for emphasis
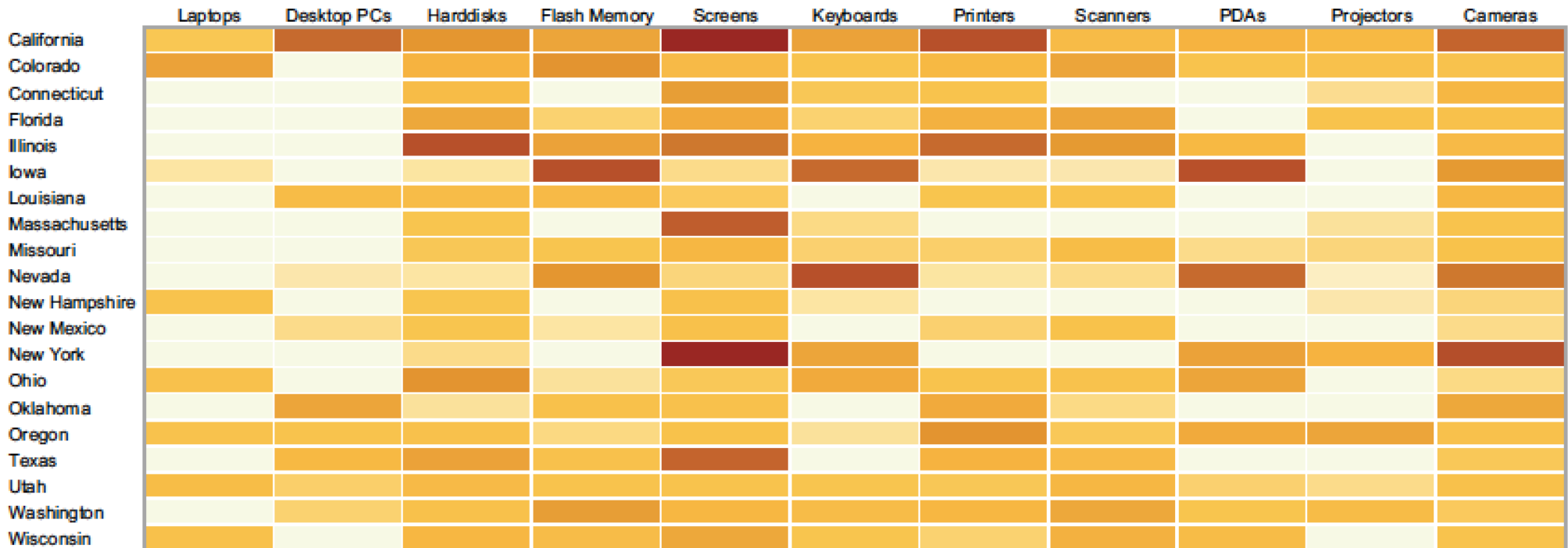- Different shades of gray are colors too, not so?

Market Share

Company D	52.44%
Our Company	13.46%
Company A	13.03%
Company C	8.85%
Company F	5.37%
Company E	4.95%
All Others	1.90%

Single/Related Hues for Quatitative Data

- Which set of color patches have a natural ordering?

Example Use of Ordered Hues

High performers stand out!

Light or Gray Colors for Non-Data

Grid lines thin and light

Employee Responses Regarding Company Goals

Remember the Colorblind

- Test on B/W

- Most common is red-green
  - Consider continuum red-blue instead of red-green
  - Test yourself at http://colorvisiontesting.com/

Bad Color – Why?

1970
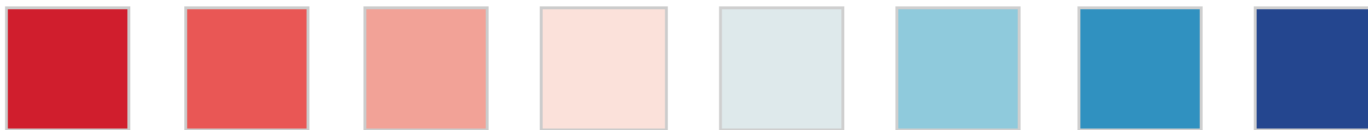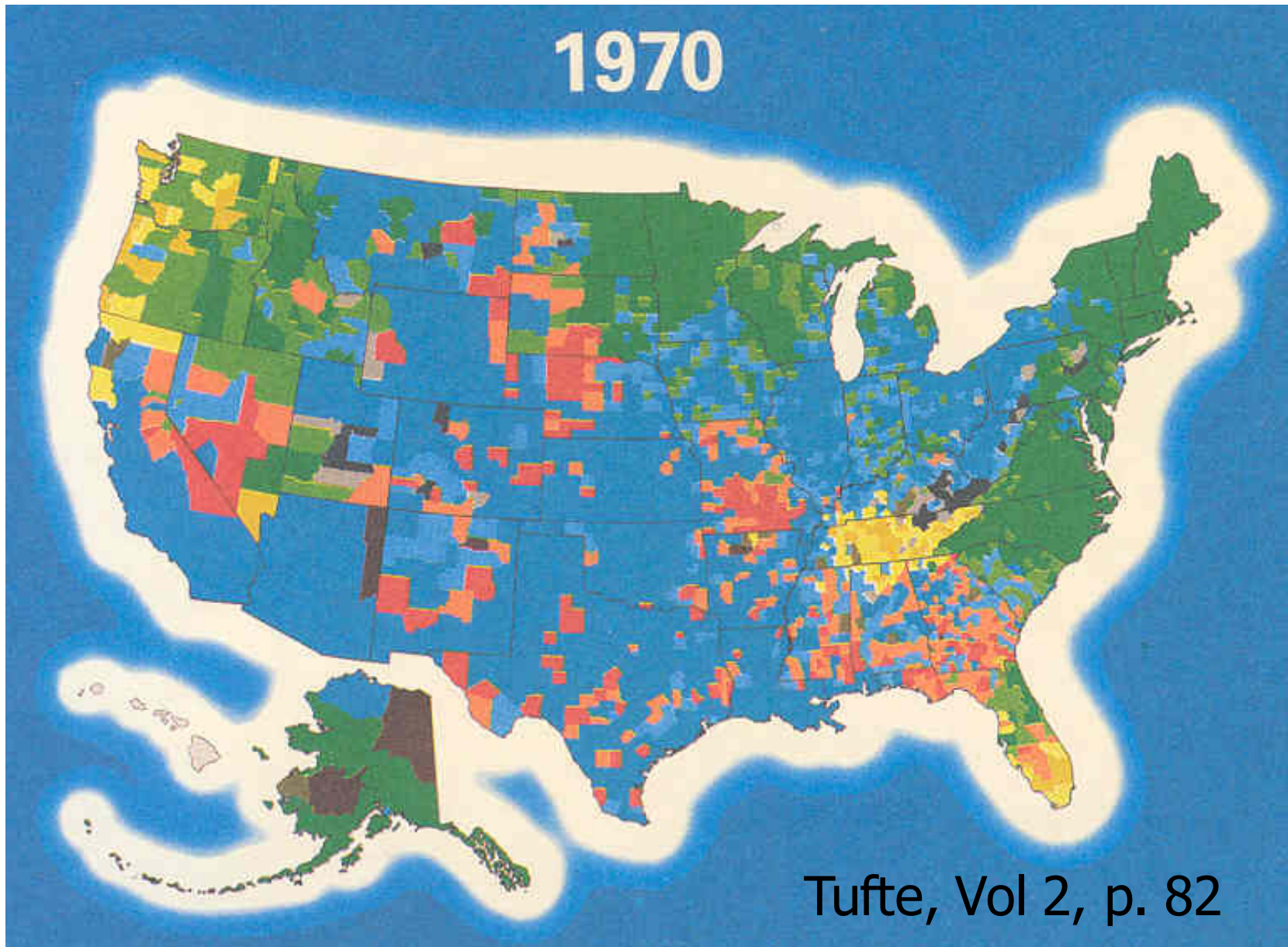
Tufte, Vol 2, p. 82

Bad Color – Why?
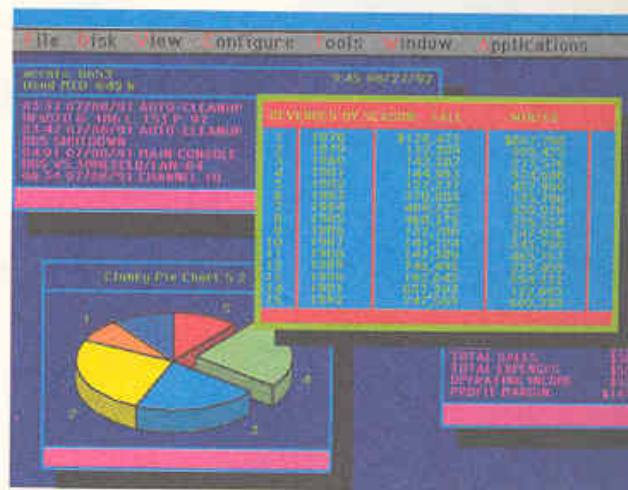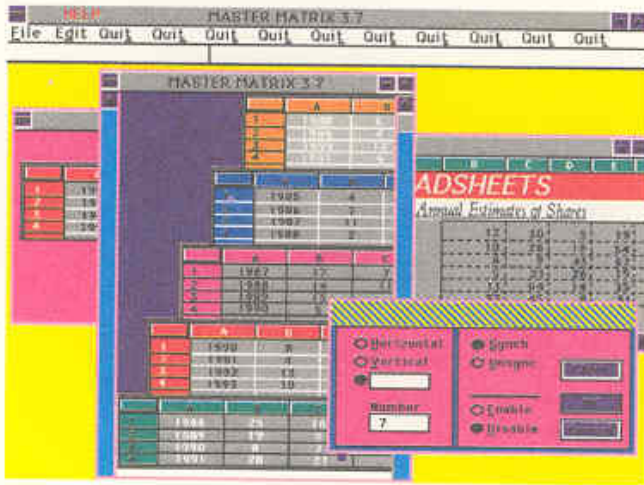
Tufte, Vol 2, p. 88

Bad Color – Why?

Bad Color? Bloomberg Financial Display

- Is black background OK?
  - Legacy of older systems
  - 8-hour use

CS 4460

Good Use of Color?

Sales by Person, Qupota vs. Actual

Good Use of Color?
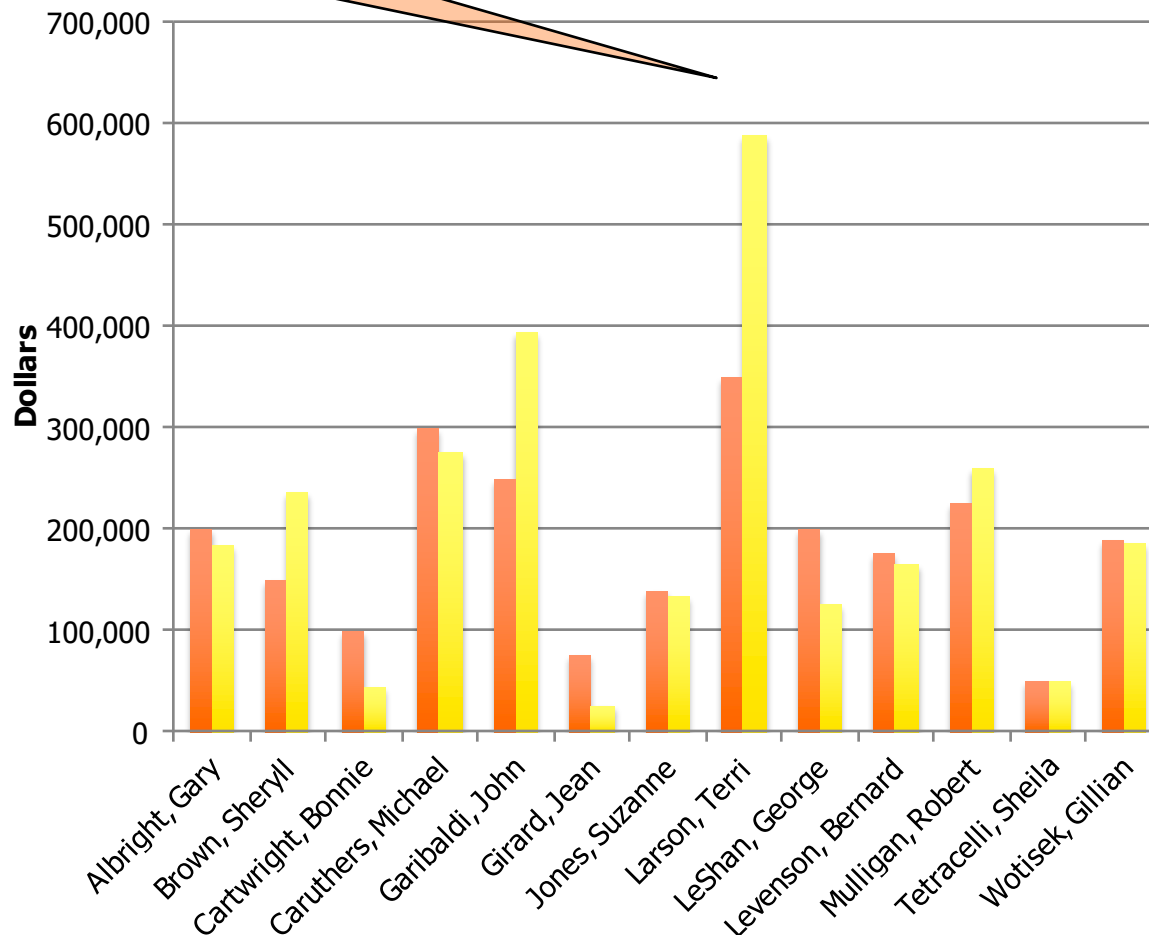
Better Use of Color

Sales by Person, Quota vs. Actual
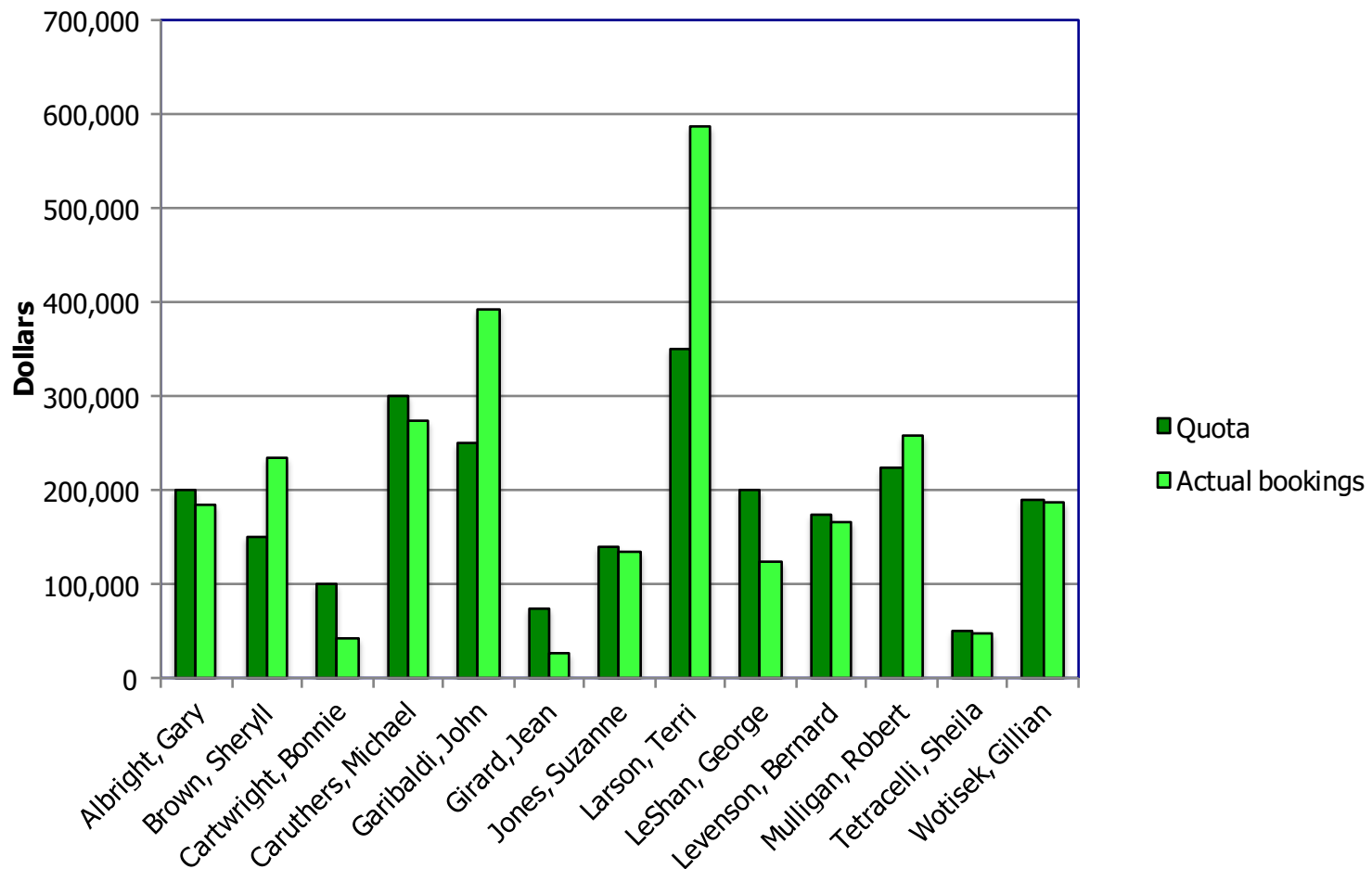
Even Better Use of Color?
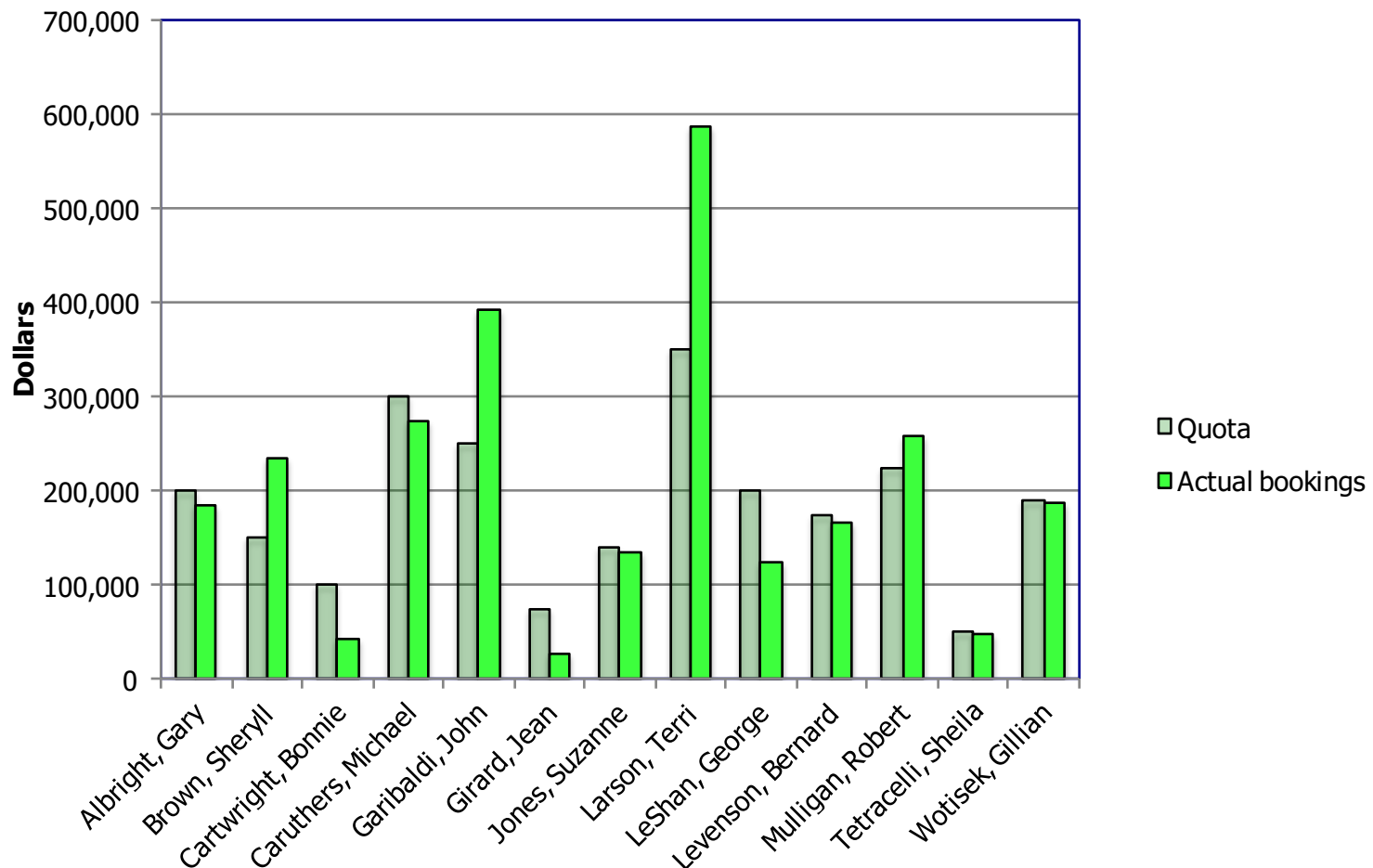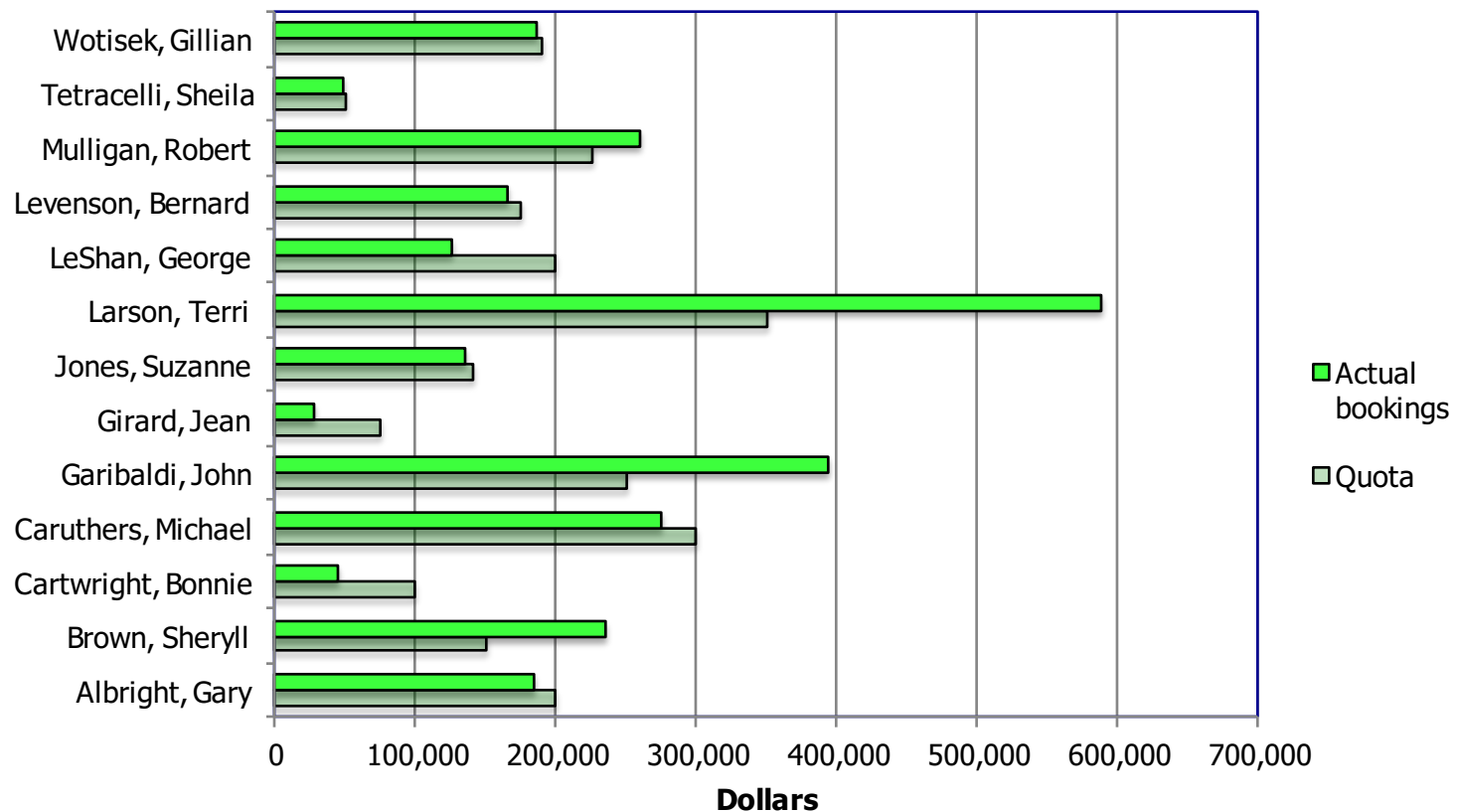

Sales by Person, Quota vs. Actual

Horizontal Bars Better with Long Labels

Sales by Person, Quota vs. Actual

Design Principle Summary

- Maximize data-ink ratio
- Content is king (less is more)
- Careful with 3D
- Avoid "chartjunk"
- Use multi-functioning graphical elements
- Be careful with color

The End

- Go Forth and Design Great Visualizations!
  - With Graphical Integrity
  - With Aesthetic Excellence

- Apply these ideas to your project!
  - Review your project's visualizations against each of the principles

    Discuss with others

    Review when meet with Tas and during work sessions