

Deceptive Statistical Graphics in the News and in Science

Jürgen Symanzik

Utah State University

Logan, UT, USA

symanzik@math.usu.edu

<http://www.math.usu.edu/~symanzik>

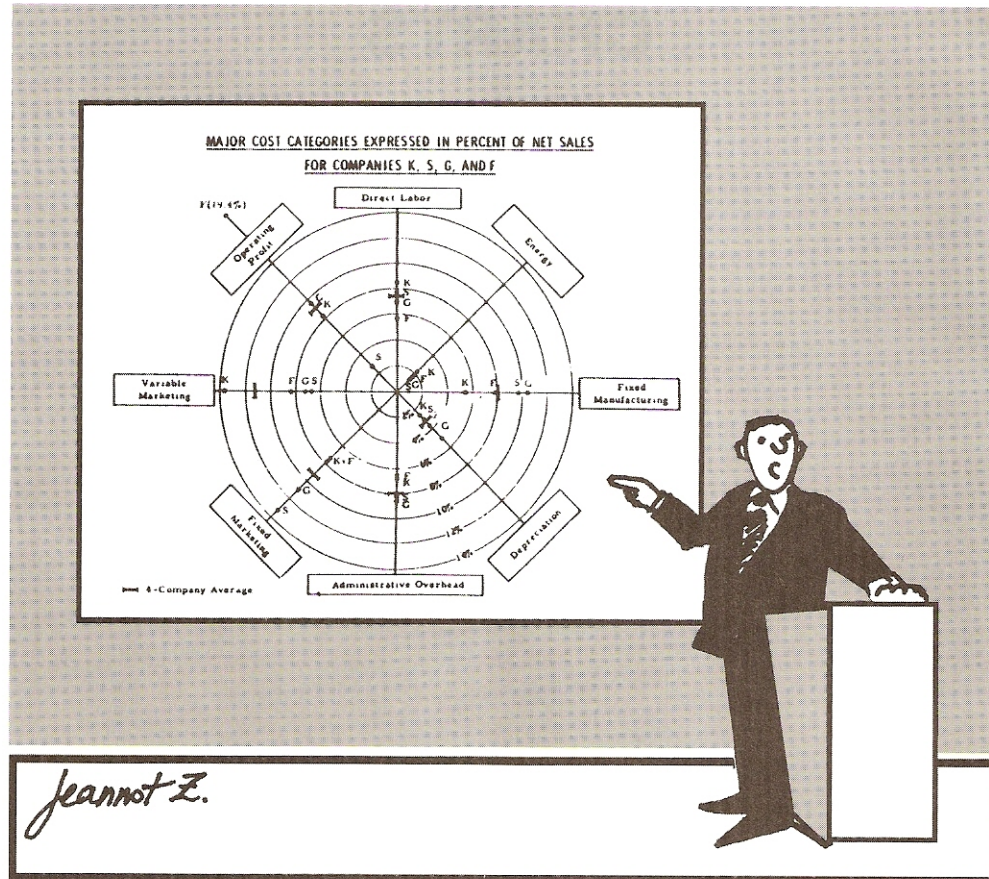
CRC 649 "Economic Risk" Conference in Motzen

July 9, 2016



Contents

- Motivation
- “How to Display Data Badly”
- Examples from News and Science
- Possible Explanations for Misleading Graphics
- Sources for Constructing Better Graphics



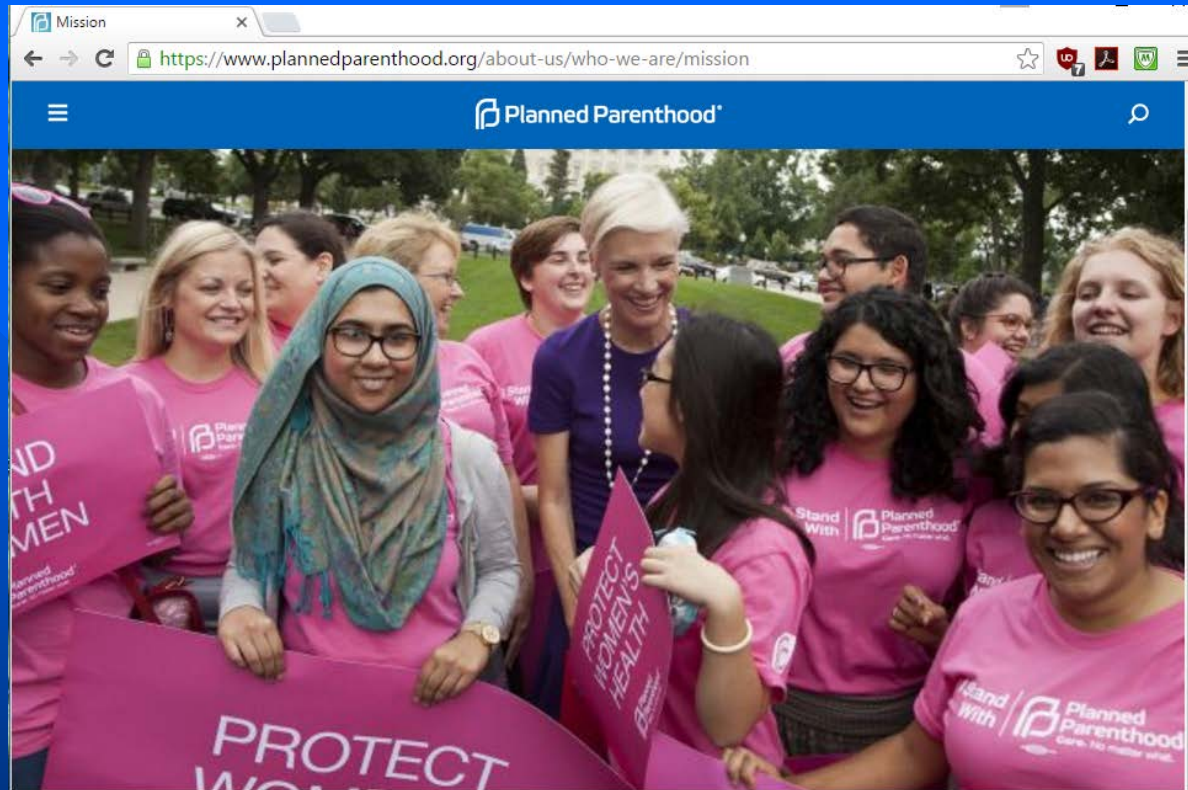
“What do you mean, what does it mean?”

From: Zelazny, G. (2001), Say it with Charts: The Executive's Guide to Visual Communication (Fourth Edition), McGraw-Hill, New York, NY.

Motivation

- As an expert in statistical graphics, it hurts to see bad graphics in the news, on the web, and produced by students.
- Even worse are bad graphics in science (including those from articles in *Computational Statistics*).
- As chair of a task force of the *Statistical Graphics Section of the American Statistical Association (ASA)*, we reevaluate the winning posters of the annual ASA poster competition for children from kindergarten to grade 12 – and see many bad graphics among the winners.

Planned Parenthood Hearing (1)



Planned Parenthood Federation of America
Mission Statement: A Reason for Being

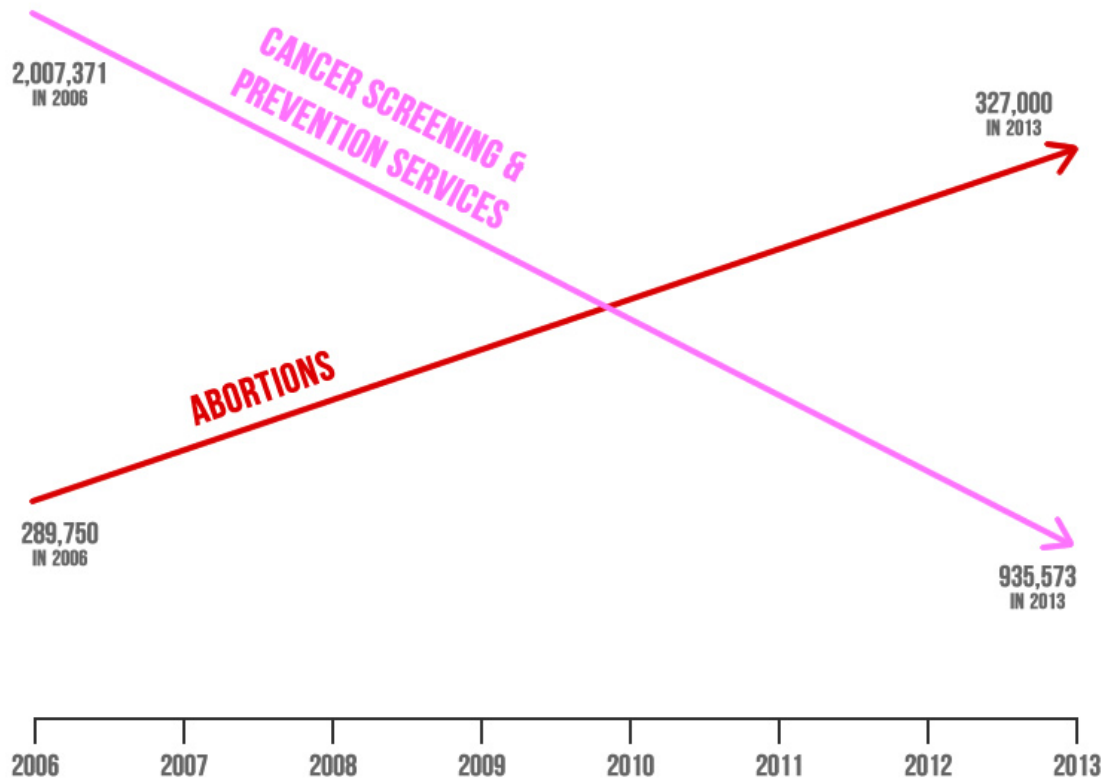
Planned Parenthood believes in the fundamental right of each individual, throughout the world, to manage his or her fertility, regardless of the individual's income, marital status, race, ethnicity, sexual orientation, age, national origin, or residence. We believe that respect and value for diversity in all aspects of our organization are essential to our well-being. We believe that reproductive self-determination must be voluntary and preserve the individual's right to privacy. We further believe that such self-determination will contribute to an enhancement of the quality of life and strong family relationships.

Based on these beliefs, and reflecting the diverse communities within which we operate, the mission of Planned Parenthood is

- to provide comprehensive reproductive and complementary health care services in settings which preserve and protect the essential privacy and rights of each individual
- to advocate public policies which guarantee these rights and ensure access to such services
- to provide educational programs which enhance understanding of individual and societal implications of human sexuality
- to promote research and the advancement of technology in reproductive health care and encourage understanding of their inherent bioethical, behavioral, and social implications

Planned Parenthood Hearing (2)

PLANNED PARENTHOOD FEDERATION OF AMERICA: ABORTIONS UP — LIFE-SAVING PROCEDURES DOWN



SOURCE: AMERICANS UNITED FOR LIFE

Planned Parenthood Hearing (3)



In 2006, Planned Parenthood performed more prevention services and cancer screenings than abortions, but in 2013, there were more abortions.

— *Jason Chaffetz on Tuesday, September 29th, 2015 in a Congressional hearing.*

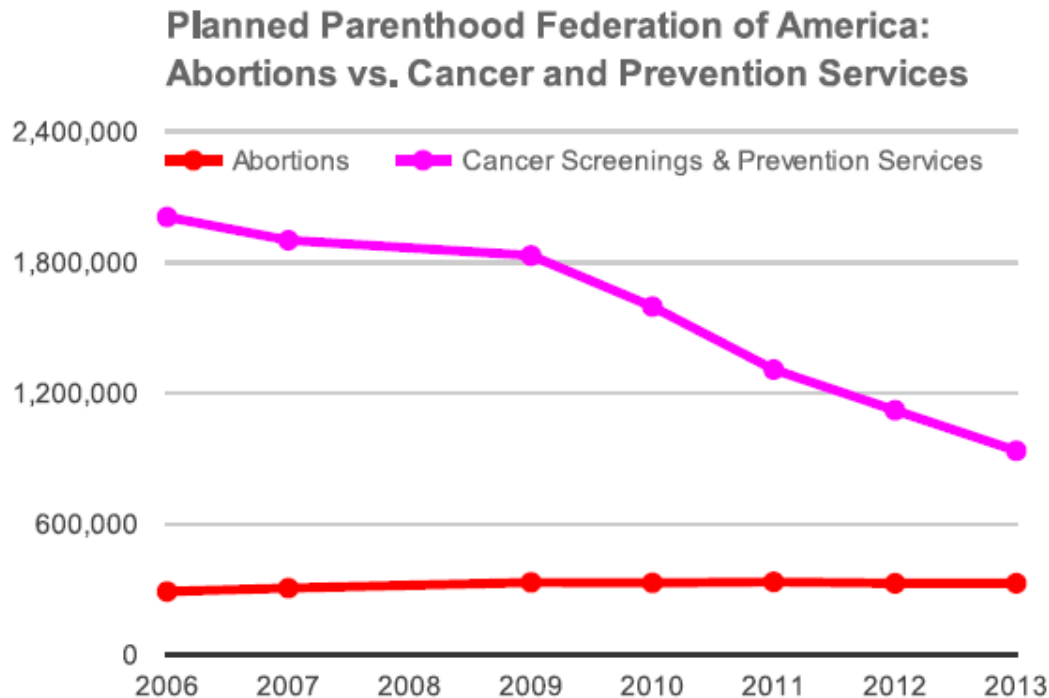
Chart shown at Planned Parenthood hearing is misleading and 'ethically wrong'

By Linda Qiu on Thursday, October 1st, 2015 at 5:08 p.m.

Planned Parenthood Hearing (4)

Corrected charts

With guidance from the experts, we compiled the number of abortions and cancer screenings/prevention services from Planned Parenthood's annual reports from 2006 to 2013 (we could not find a report from 2008). Here's what the chart should look like:



Planned Parenthood Hearing (5)

Our ruling

At the hearing, Chaffetz presented a chart that showed the number of abortions at Planned Parenthood rising higher than the number of preventive services and cancer screenings between 2006 and 2013.

But that's inaccurate, disputed by the chart's own, hard to read numerical labels. In fact, there were three times as many cancer screenings and prevention services as abortions in 2013. Experts in data presentation said this was an egregious example of using a chart to mislead.

We rate the claim Pants on Fire.

Obamacare (1)



ObamaCare Facts: An Inde X

obamacarefacts.com

OBAMACARE FACTS

ObamaCare Facts: An Independent Site For ACA Advice



OBAMACARE ALTERNATIVE

Click to Enroll

Call Now: **1-800-508-6754**

Ads by +HealthNetwork

Short-Term Health Insurance Obamacare Alternative Obamacare Health Plans

What is The Affordable Care Act?

The Affordable Care Act (ACA), officially called [The Patient Protection and Affordable Care Act \(PPACA\)](#) and sometimes called ObamaCare, is a US law that reforms both the healthcare and health insurance industries in America.

The PPACA was signed into law by President Barack Obama on March 23, 2010 and upheld by the Supreme Court on June 28, 2012. Despite many repeal attempts, the ACA is still the “law of the land today,” that means lower costs on coverage, new benefits, and a few new taxes. We explain below.

What Does the Affordable Care Act Do?

The law aims to increase the quality, availability, and affordability of private and public health insurance to over 44 million uninsured Americans through its many provisions, which include new regulations, taxes, mandates, and subsidies. The law also works toward curbing the growth in healthcare spending in the US, which has been rising at an [unsustainable rate](#).

Obamacare (2)

OBAMACARE ENROLLMENT



AS OF
MARCH 27

MARCH 31
GOAL

SOURCE: HHS



mediamatters.org

IE IN SECOND PLACE WITH \$26.5 MIL, WHILE "MUPP DOW FUT 16,325.00

Obamacare (3)

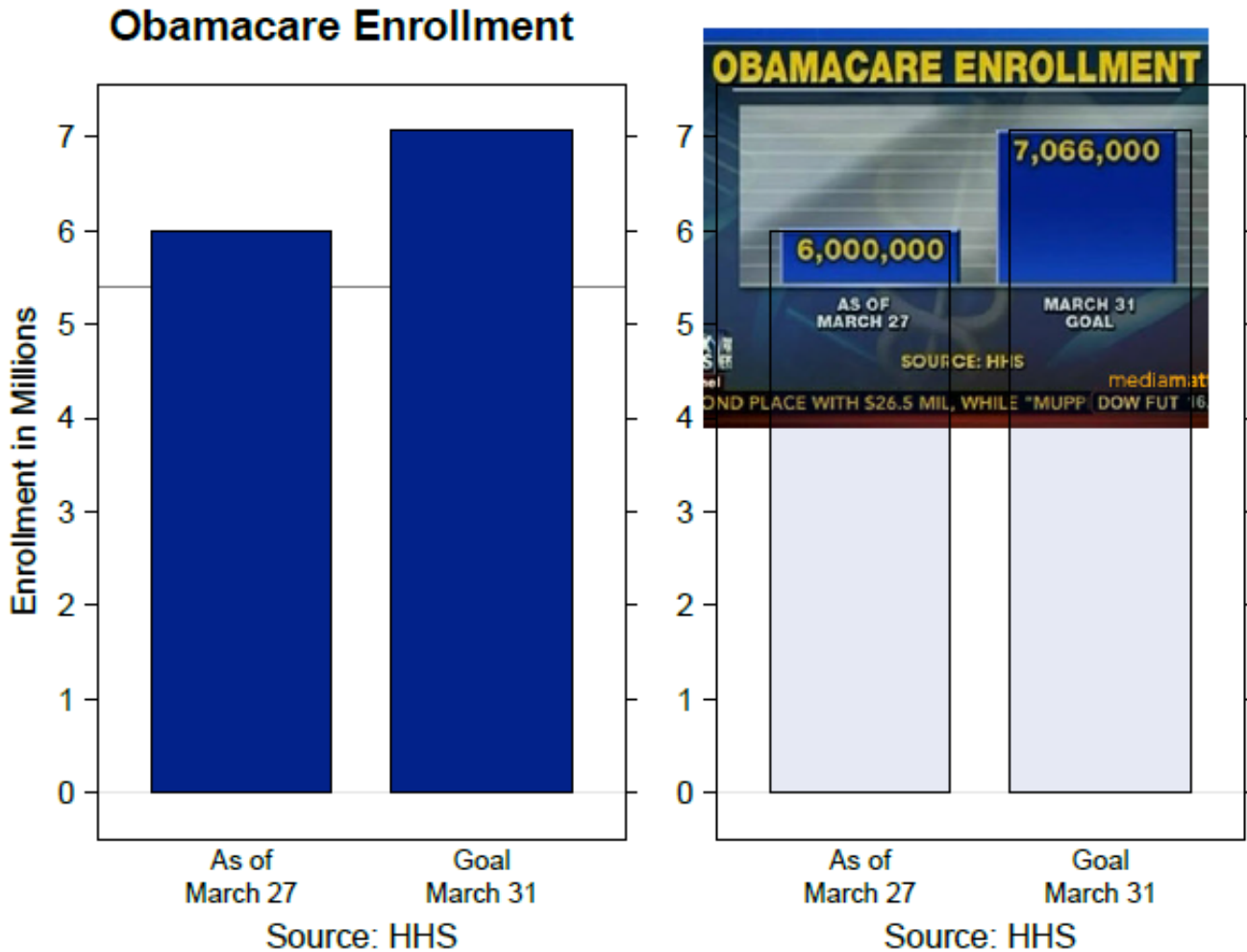


Figure 10: Comparison of misleading graph with accurate graph. The left panel shows an accurately drawn bar chart, with origin at 0, of the same data. We now see that the bar

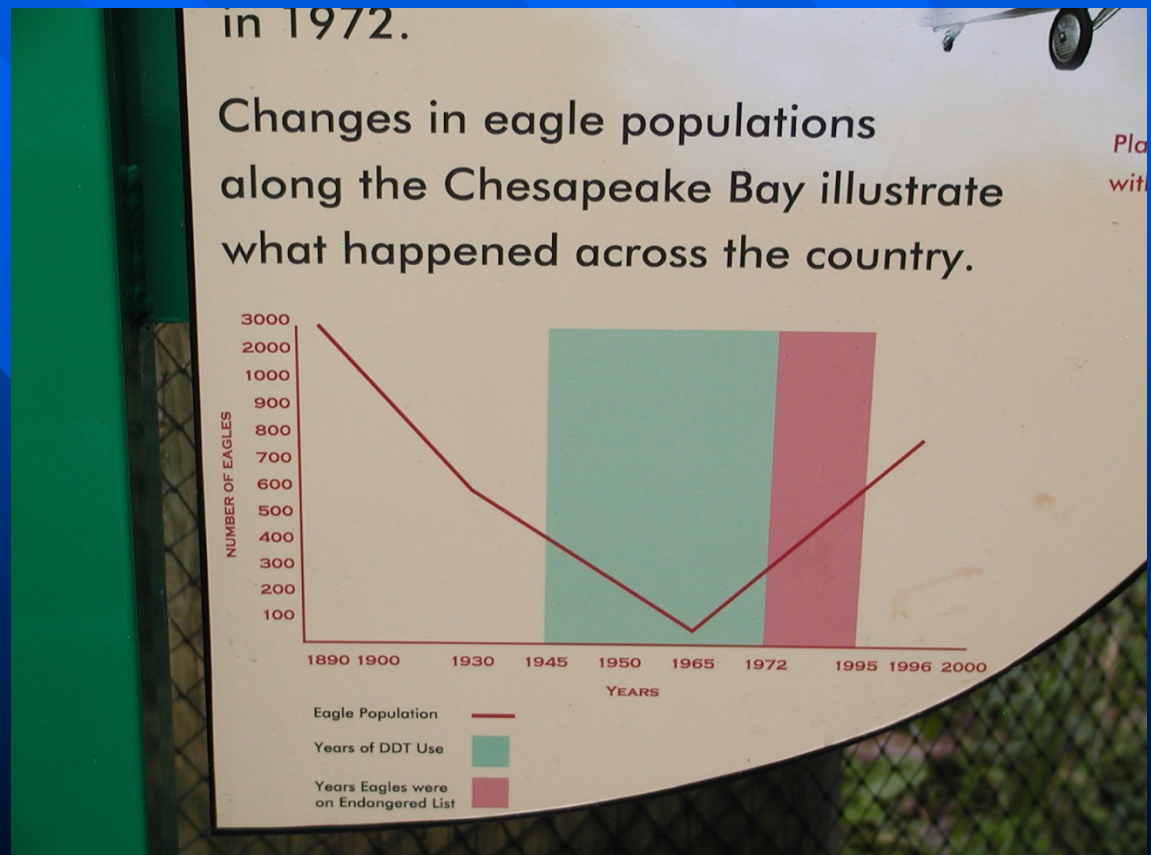
From: Heiberger, R. M.,
Robbins, N. B.,
Symanzik, J. (2014).
“Statistical Graphics
Recommendations for
the ASA/NCTM Annual
Poster Competition and
Project Competition”,
JSM Proceedings,
American Statistical
Association, Alexandria,
VA.

“How to Display Data Badly”

- From: Wainer, H. (1997), *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*, Copernicus/Springer, New York, NY:
- “The aim of good data graphics is to display data accurately and clearly. [...]
- Thus, if we wish to display data badly, we have three avenues to follow.
 - A. Don't show much data.
 - B. Show the data inaccurately.
 - C. Obfuscate the data.” [i.e., show the data unclearly]

A. Don't show much data

- Rule 1: Show as little data as possible (minimize the data density).



From: Zoo, Indianapolis, IN.
Personal Photograph, July 11, 2008.

A. Don't show much data

- Rule 2: Hide what data you do show (minimize the data/ink ratio).

Hispanic enrollment rises, other groups falling

By Chris Campbell
staff writer

While enrollment numbers for Hispanics are increasing, enrollment numbers for other races and ethnicities has decreased over the last four years, according to data points graphed from USU's Office of Analysis, Assessment and Accreditation.

Students said possible reasons for the upward trend in Hispanic enrollment include an increase in the Hispanic population in the U.S. as a whole and an increased determination in the Latino culture to succeed.

Leo Torres-Reyes, a freshman majoring

in biology and member of Latino Student Union, or LSU, said many Hispanics are starting to understand college is an essential way to be successful.

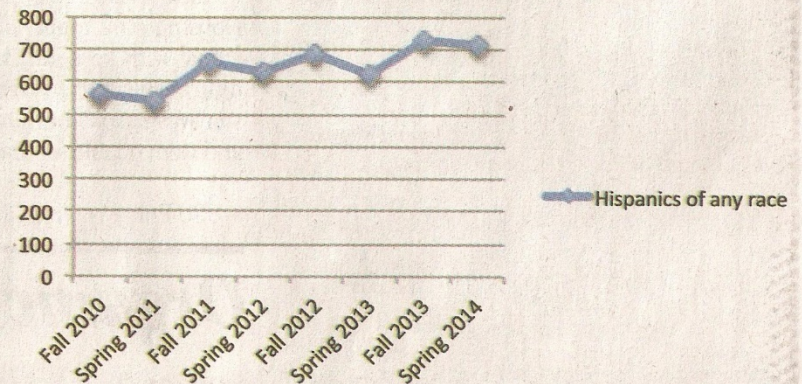
"Our parents are coming into this country for us to have a better future," Torres-Reyes said. "And I think a lot of Hispanics are starting to realize that college is what's going to get them further into the future."

Sandra Martinez, a sophomore majoring in sociology and the president of LSU, said the Latino population is growing in the U.S. in general as Mexicans immigrate to get a better job and education.

"Not that all the other ethnicities and

See ENROLL, Page 3

Hispanics of any race



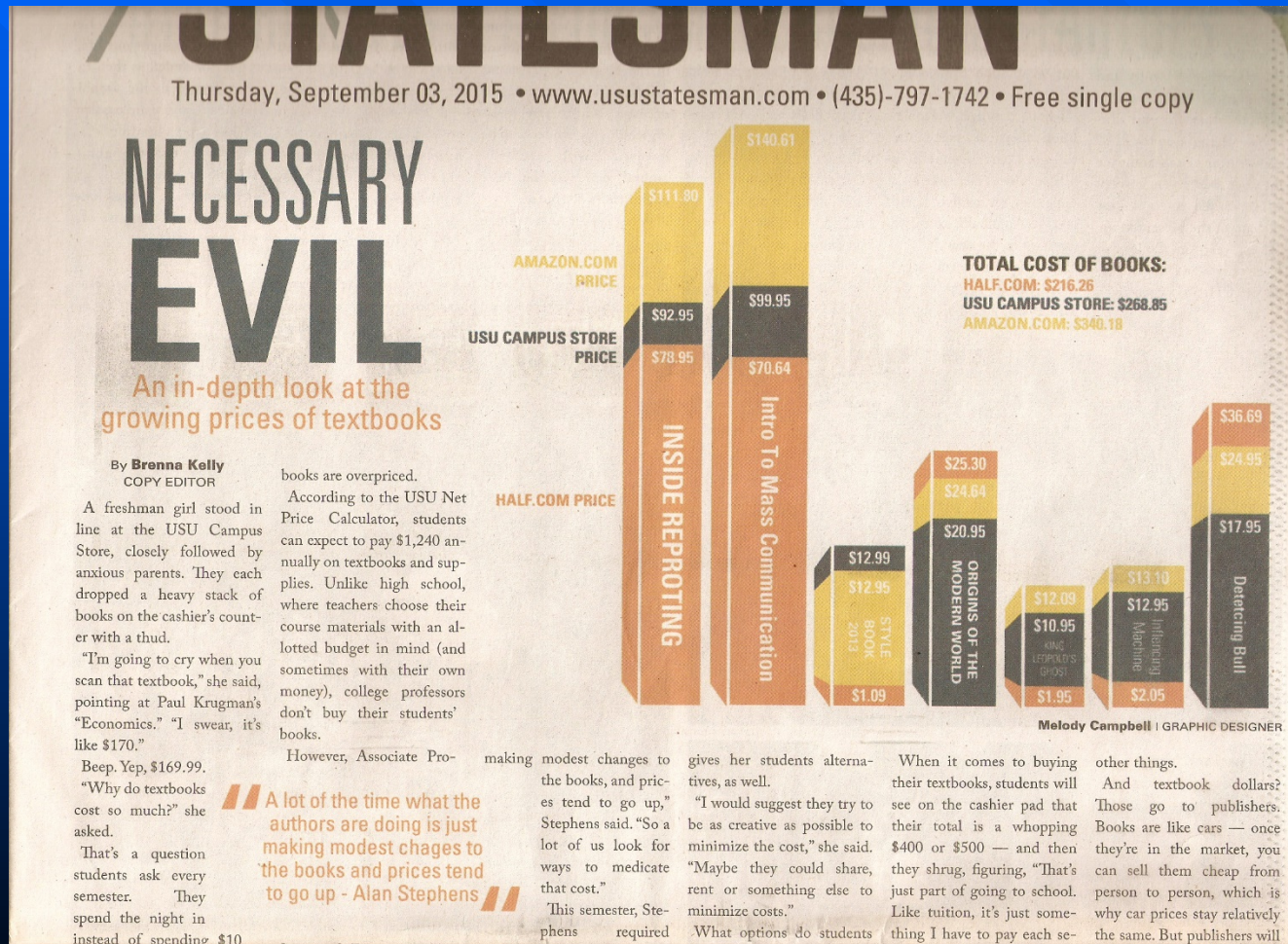
Info courtesy USU Office of Analysis, Assessment and Accreditation

The Utah Statesman, Thu, 4/24/2014

Utah congressman pushes national parks bill

B. Show the data inaccurately

- Rule 3: Ignore the visual metaphor altogether.



B. Show the data inaccurately

- Rule 4: Only order matters.



From: The Utah Statesman, Online Edition, October 19, 2009

B. Show the data inaccurately

- Rule 5: Graph data out of context.

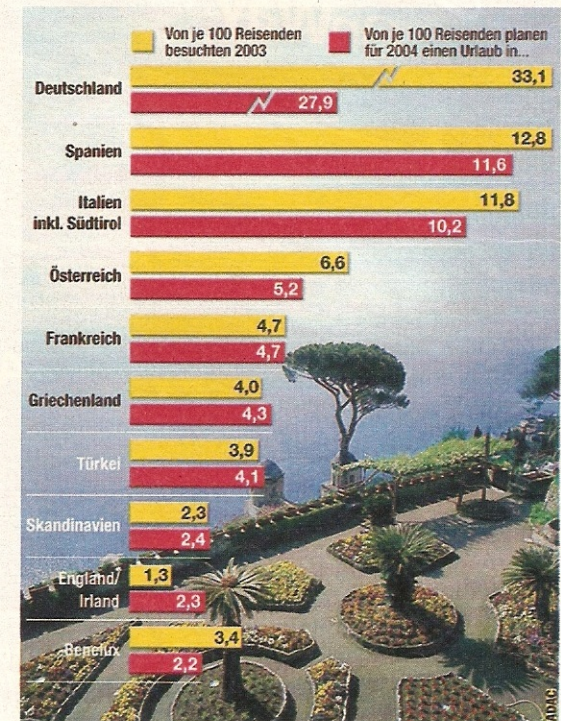


C. Obfuscate the data

- Rule 6: Change scales in mid-axis.

From: ADAC Motorwelt, April 2004.

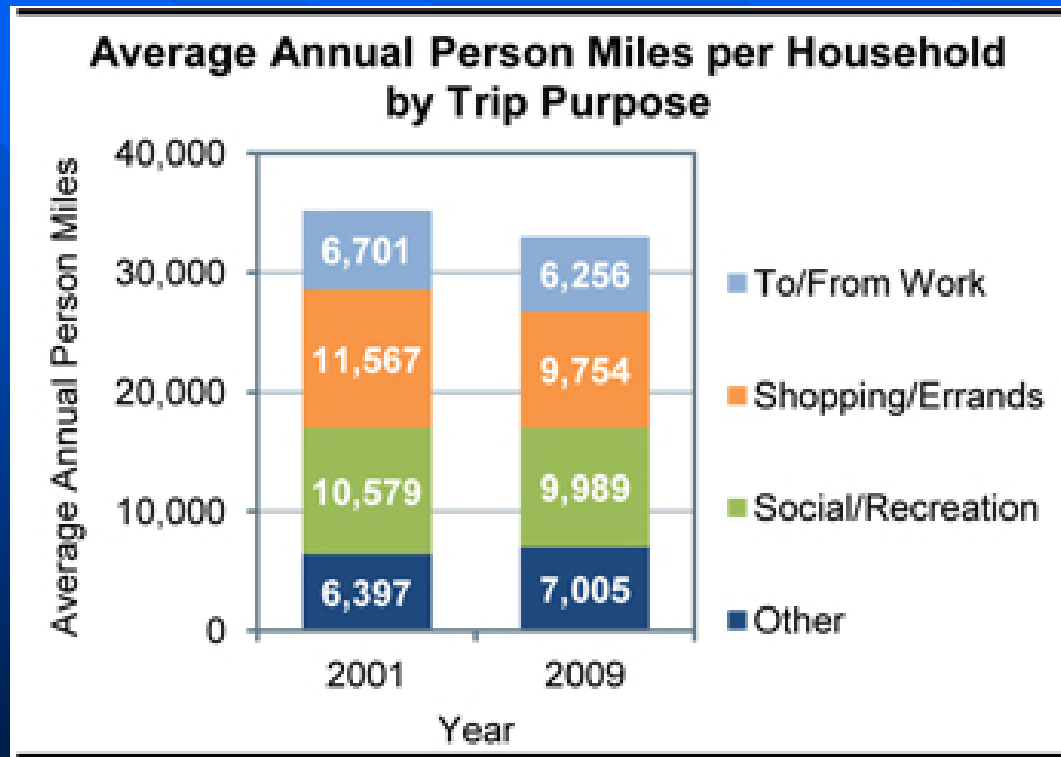
Reiseziel Spanien ganz vorn



TREND. Spanien war 2003 das wichtigste ausländische Reiseziel der Deutschen. Laut Reisemonitor, einer Umfrage des ADAC-Verlages unter 4000 Mitgliedern, machten 12,8 Prozent dort Urlaub. Italien wurde mit 11,8 Prozent vom ersten Platz verdrängt. Die meisten Urlauber zogen die Heimat dem Ausland vor: 33,1 Prozent blieben in heimischen Gefilden (siehe auch S. 104).

C. Obfuscate the data

- Rule 7: Emphasize the trivial (ignore the important).



From: Federal Highway Administration, U.S. Department of Transportation
<http://www.fhwa.dot.gov/policy/2013cpr/overviews.cfm#exhibites01a>

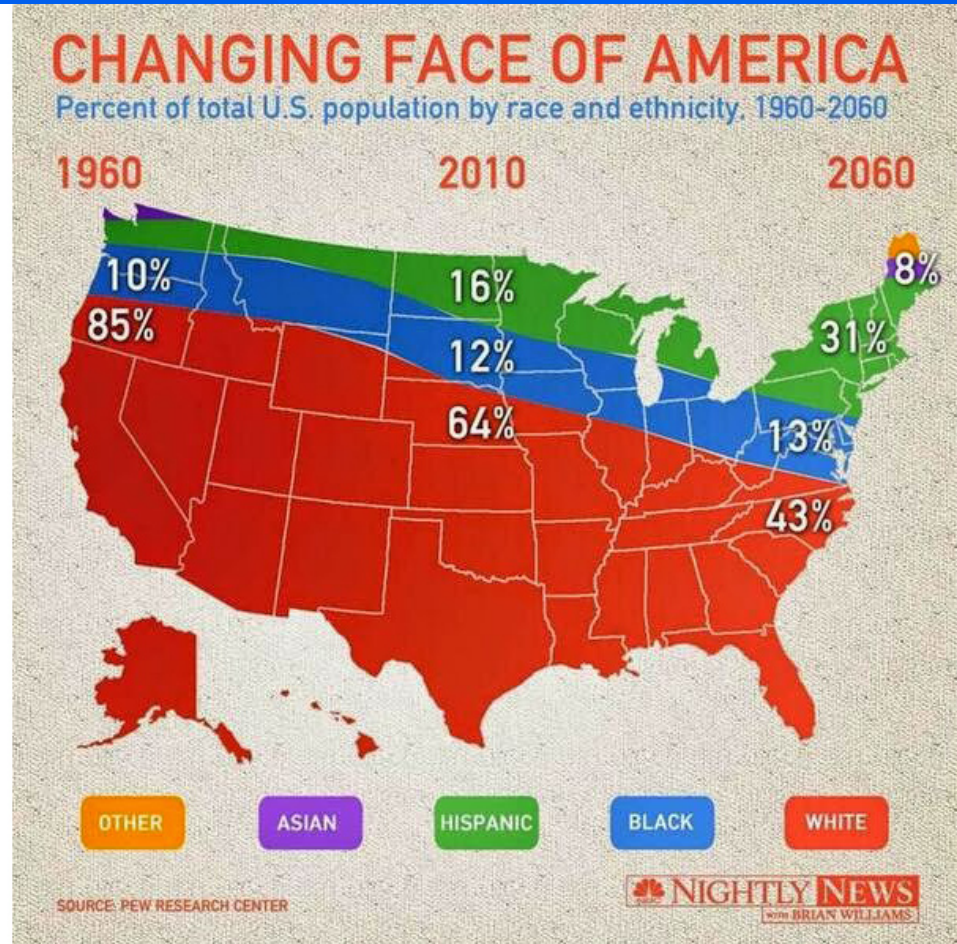
C. Obfuscate the data

- Rule 8: Jiggle the baseline.

Original: NBC Nightly News

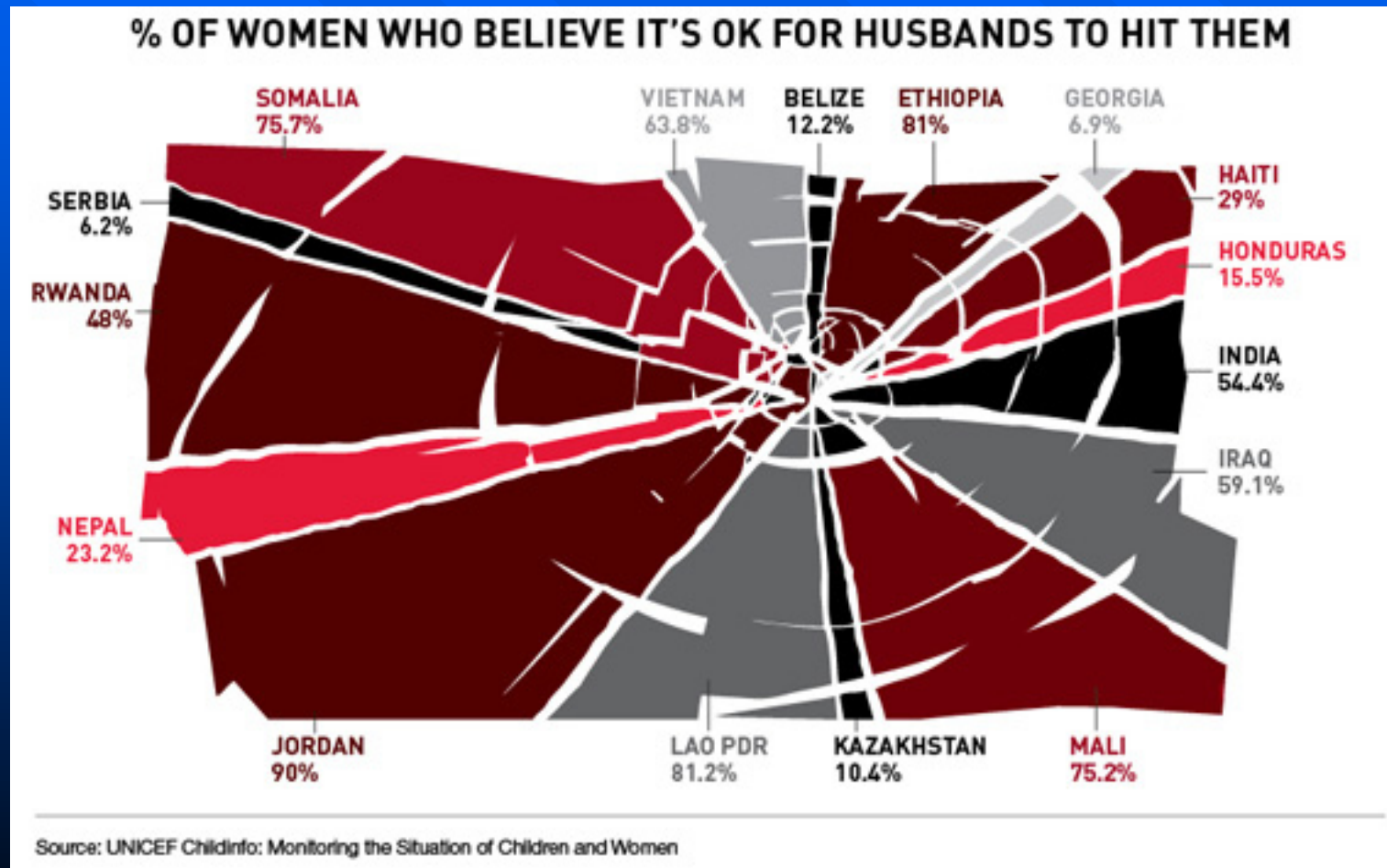
Discussed at

<http://www.fastcodesign.com/3029318/design-crime/the-worst-infographic-of-2014-so-far>



C. Obfuscate the data

- Rule 9: Alabama first!



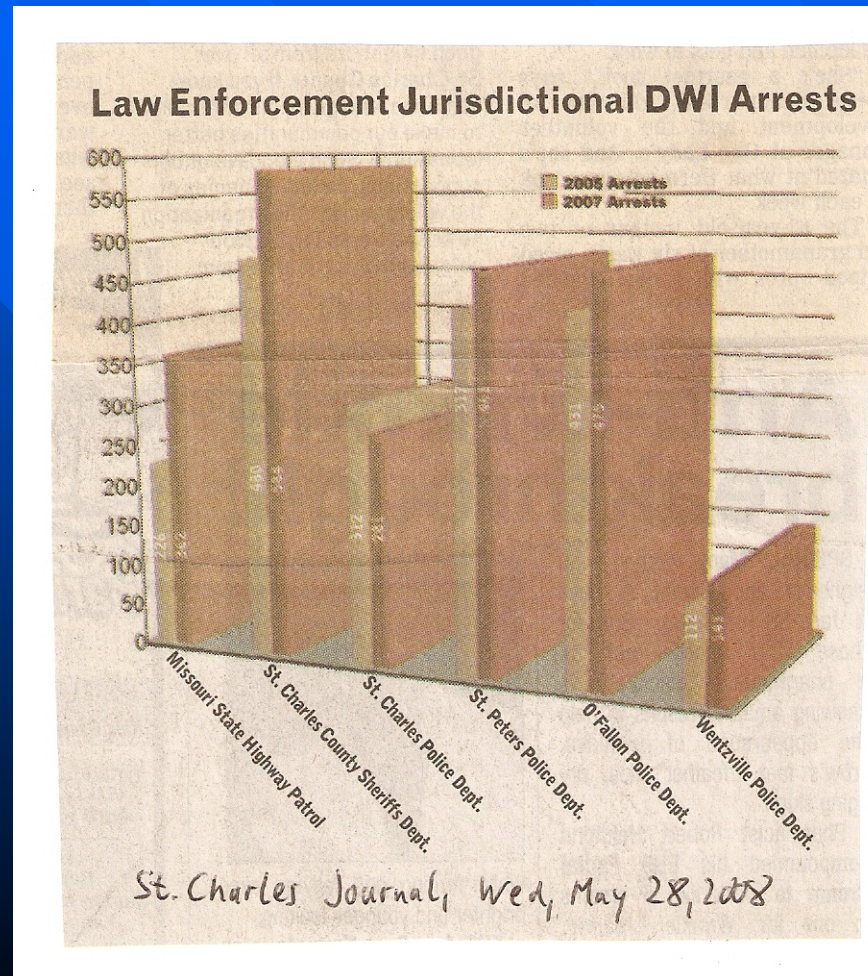
C. Obfuscate the data

- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.



C. Obfuscate the data

- Rule 11: More is murkier: (a) more decimal places and (b) more dimensions.



C. Obfuscate the data

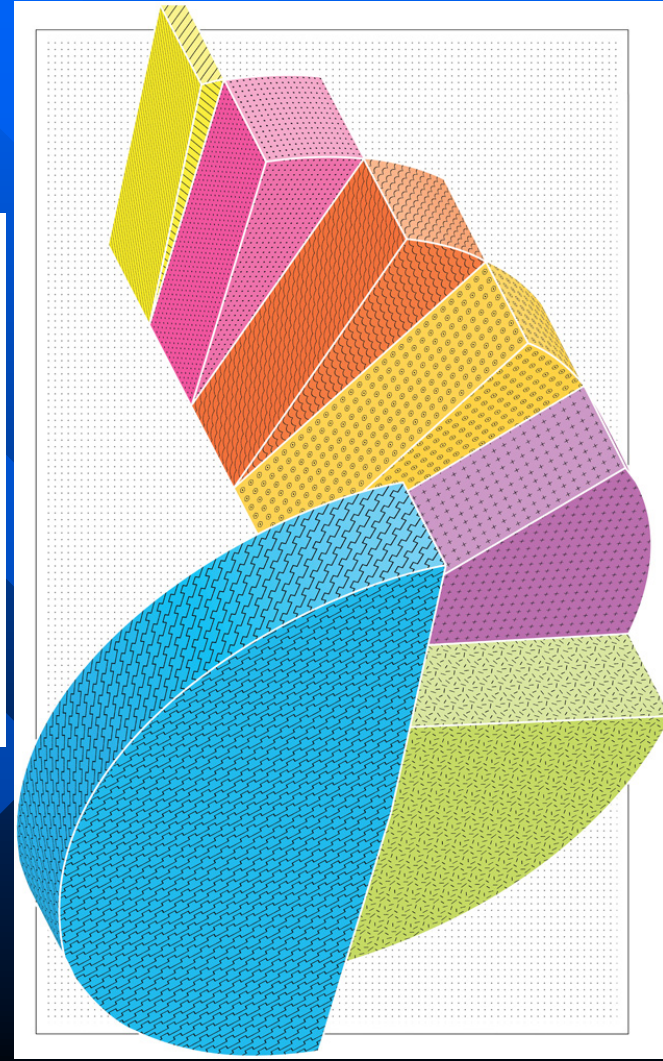
- Rule 12: If it has been done well in the past, think of a new way to do it.

JASON KEHE MAGAZINE 04.16.13 6:30 AM

WIRED'S GUIDE TO CRAFTING THE PERFECT TED TALK

IF YOU'RE INVITED to speak at TED—the annual tech, entertainment, and design conference—you say yes. TED talks have become a rite of passage for thought leaders: You're an expert in your field, and it's time to tell the world. That's a lot of pressure. But if you watch the most popular TED talks, you'll notice that there's a formula. So for you aspiring toastmasters out there, here's how to optimize your moment in the spotlight.

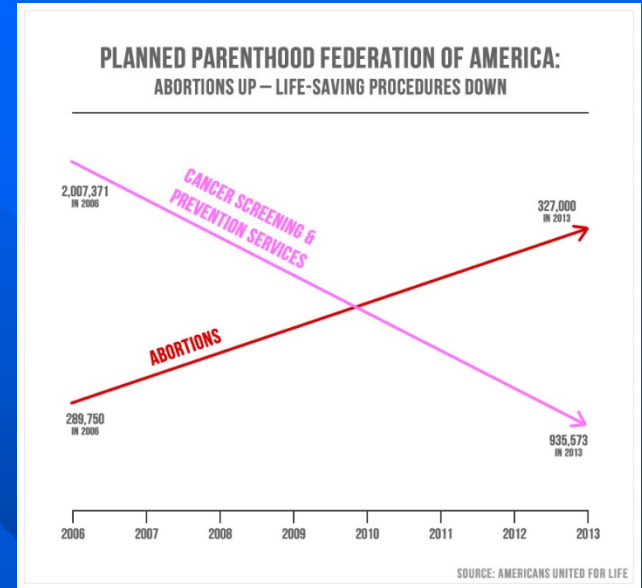
ANATOMY OF A WINNING TED TALK



From: <http://www.wired.com/2013/04/tedtalk/>

Examples from Around the World

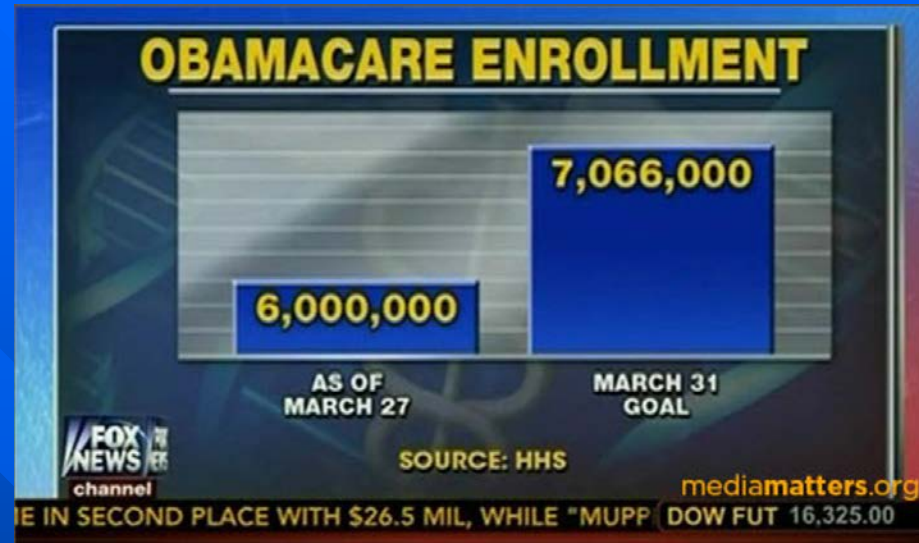
Planned Parenthood Hearing - Revisited



Rules violated:

- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.
- Rule 3: Ignore the visual metaphor altogether.
- Rule 5: Graph data out of context.

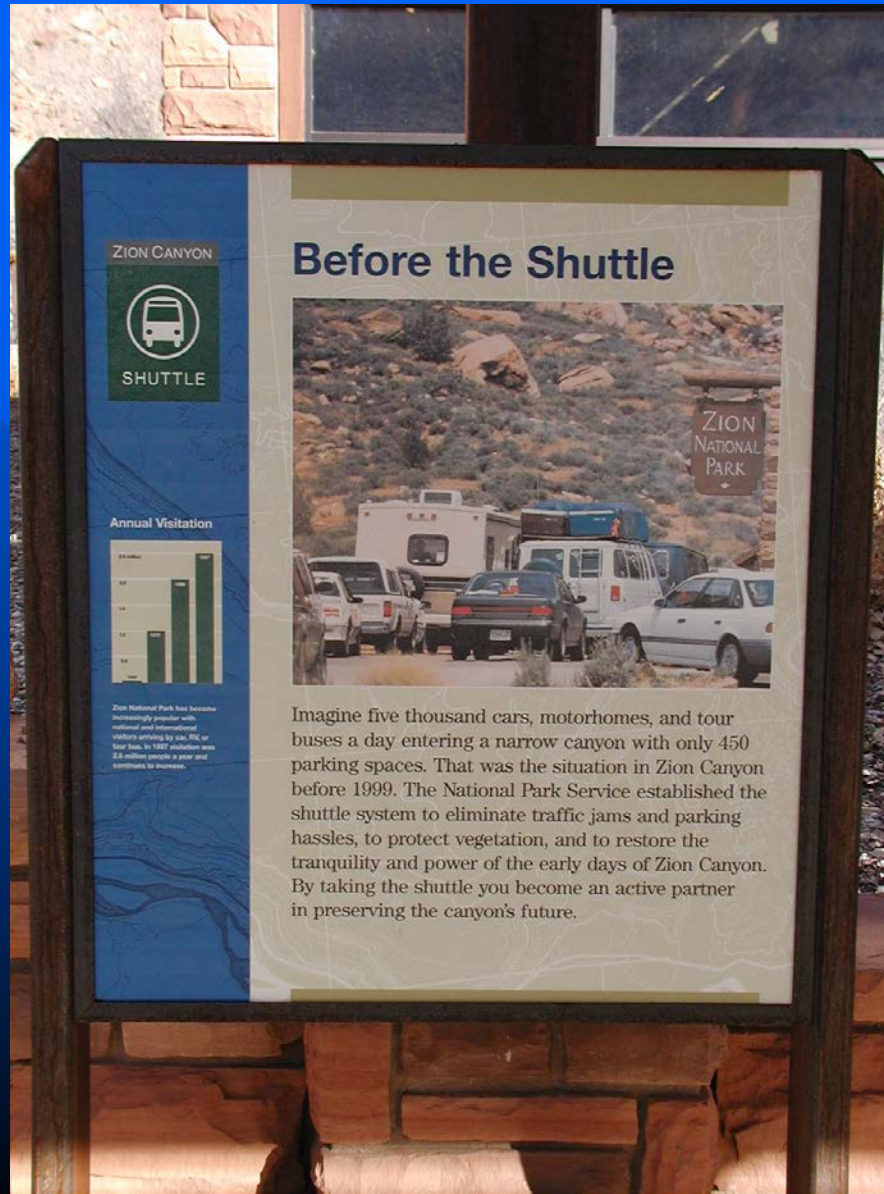
Obamacare - Revisited



Rules violated:

- Rule 4: Only Order Matters.
- Rule 1: Don't Show Much Data.
- Rule 5: Graph data out of context.
- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.

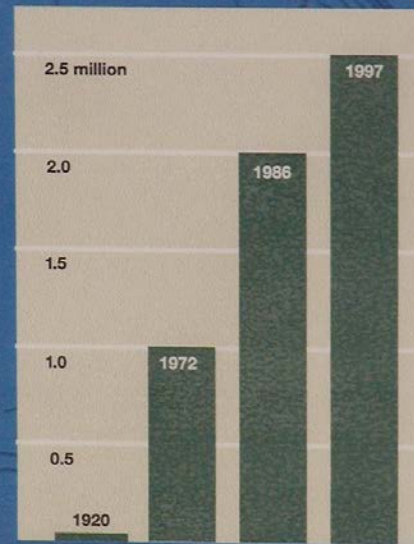
Zion National Park (1)



From: Zion National Park, UT, Shuttle Parking Lot. Personal Photograph, December 28, 2002.

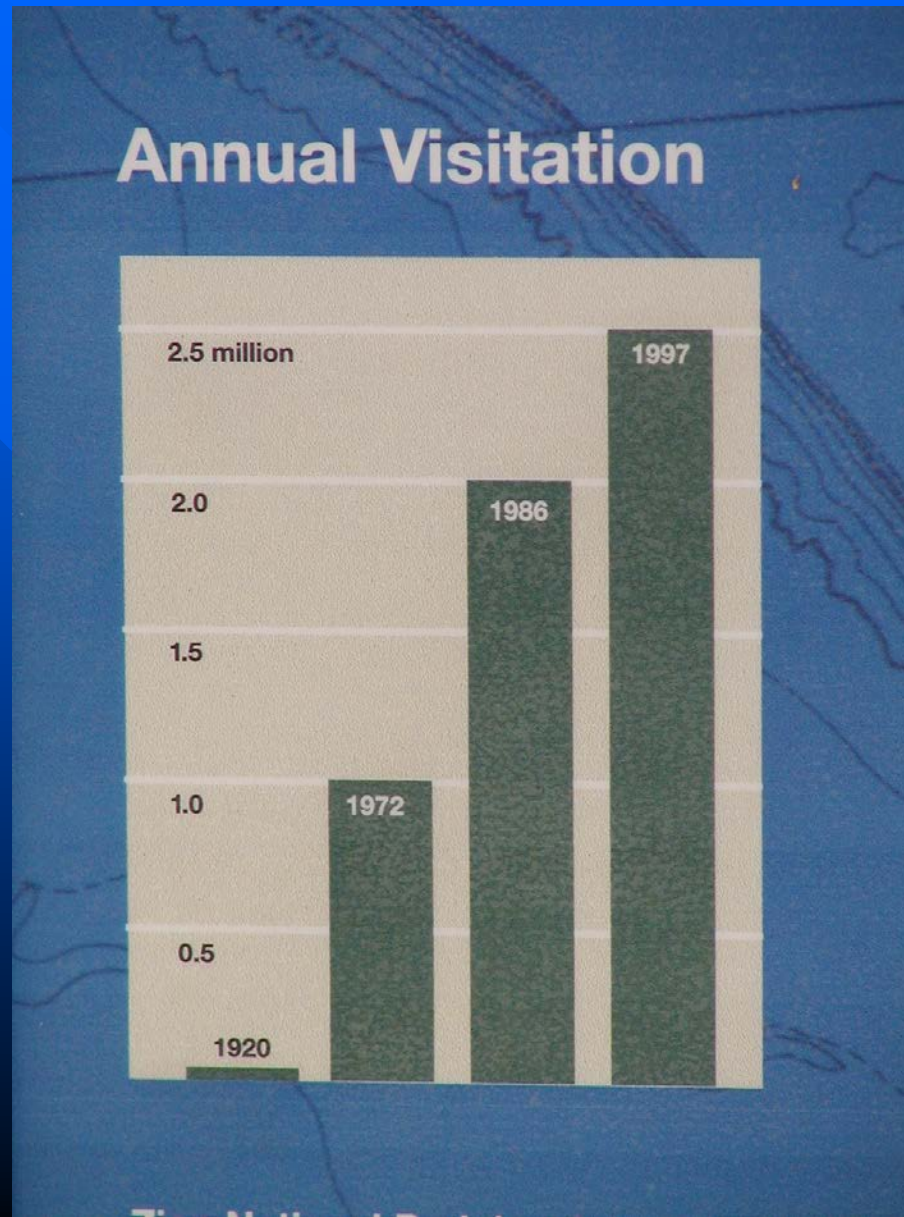
Zion National Park (2)

Annual Visitation



Zion National Park has become increasingly popular with national and international visitors arriving by car, RV, or tour bus. In 1997 visitation was 2.5 million people a year and continues to increase.

Zion National Park (3)



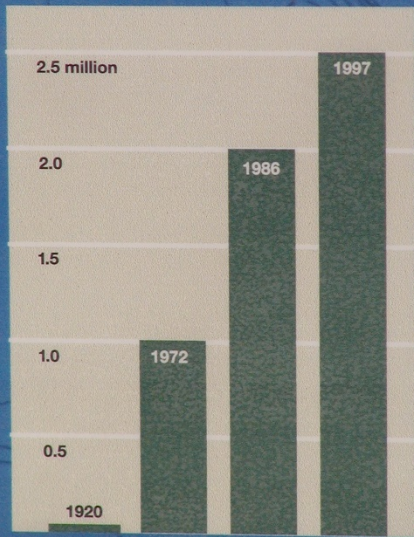
Zion National Park (4)

Rules violated:

- Rule 6: Change scales in mid-axis.
- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.
- Rule 1: Show as little data as possible (minimize the data density).
- Rule 5: Graph data out of context.

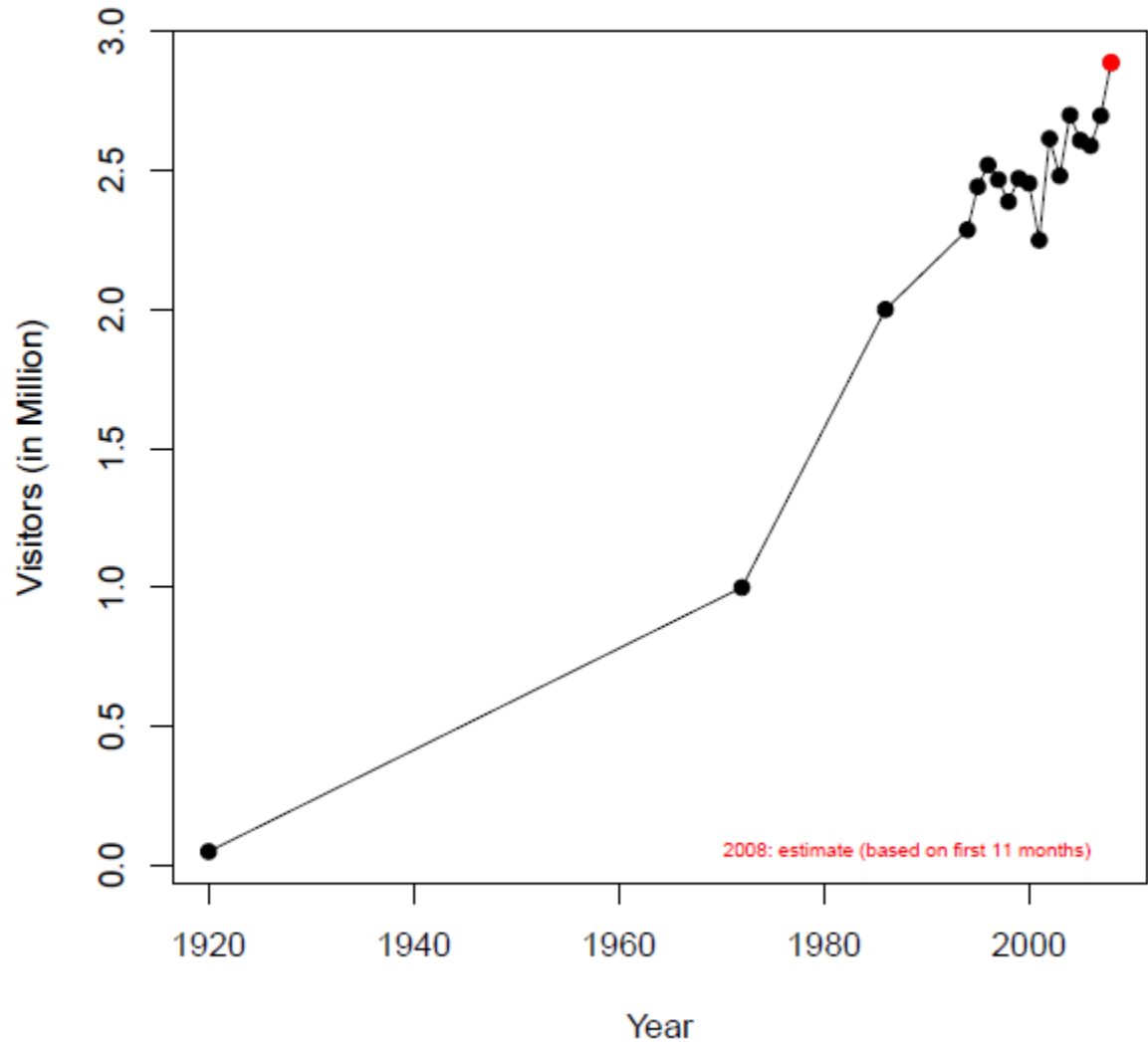
Zion National Park (5)

Annual Visitation



Improved
Graph

Annual Visitation of Zion NP



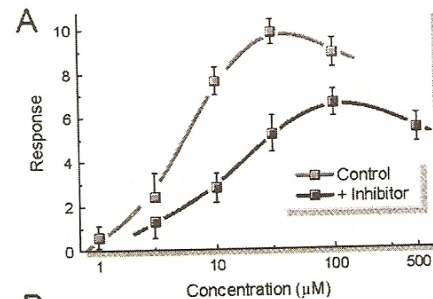
(1994-2008 data from <http://www.nps.gov/zion/parkmgmt/park-visitation-statistics.htm>)

Computational Statistics 2002 (1)

430

age, surface and contour graphs. Various types of graphs created by KyPlot are shown in Figures 3 and 4.

Almost every component of each graph can be customized through dialog boxes. Double-clicking an axis of a graph brings up a dialog box through which one can change various settings for the axis interactively. The scales of the x- and y-axes of graphs can be individually set as either linear or logarithmic. Error bars can be attached to either x- or y-values, or both, and the attributes of individual error bars can also be customized. (For example, in Figure 3A, the error bars for two data points of a line graph have been partially suppressed to avoid overlapping.) A break along an axis can be set, over a specific range and at a specific location, to indicate that a range of values has been omitted (Figure 3B).



B
Effect of low Na⁺ on high K⁺-evoked amino acid release

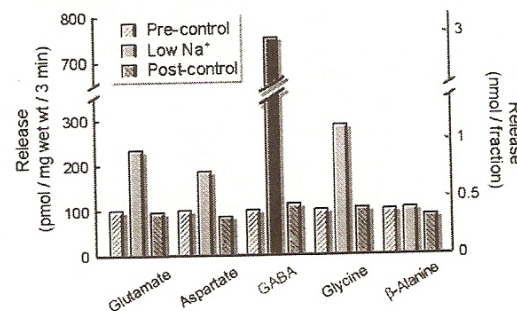
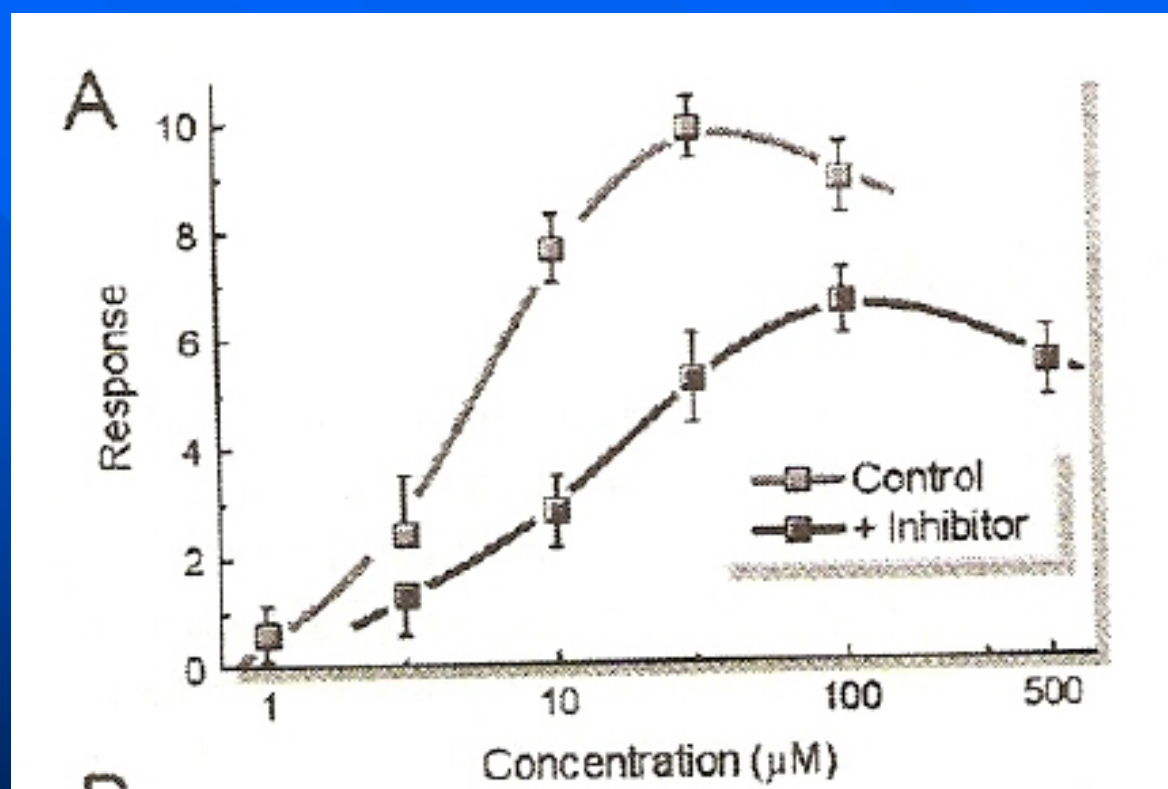


Figure 3: Line and bar graphs created with KyPlot

From: Yoshioka, K. (2002), 'KyPlot - A User-Oriented Tool for Statistical Data Analysis and Visualization', *Computational Statistics* 17(3), 425-437.

Computational Statistics 2002 (2)

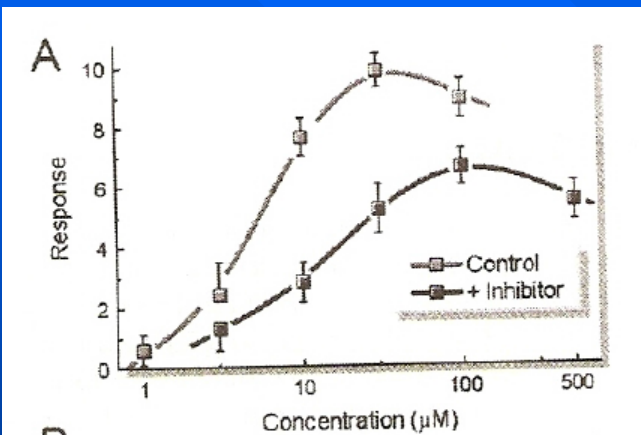


Computational Statistics 2002 (3)

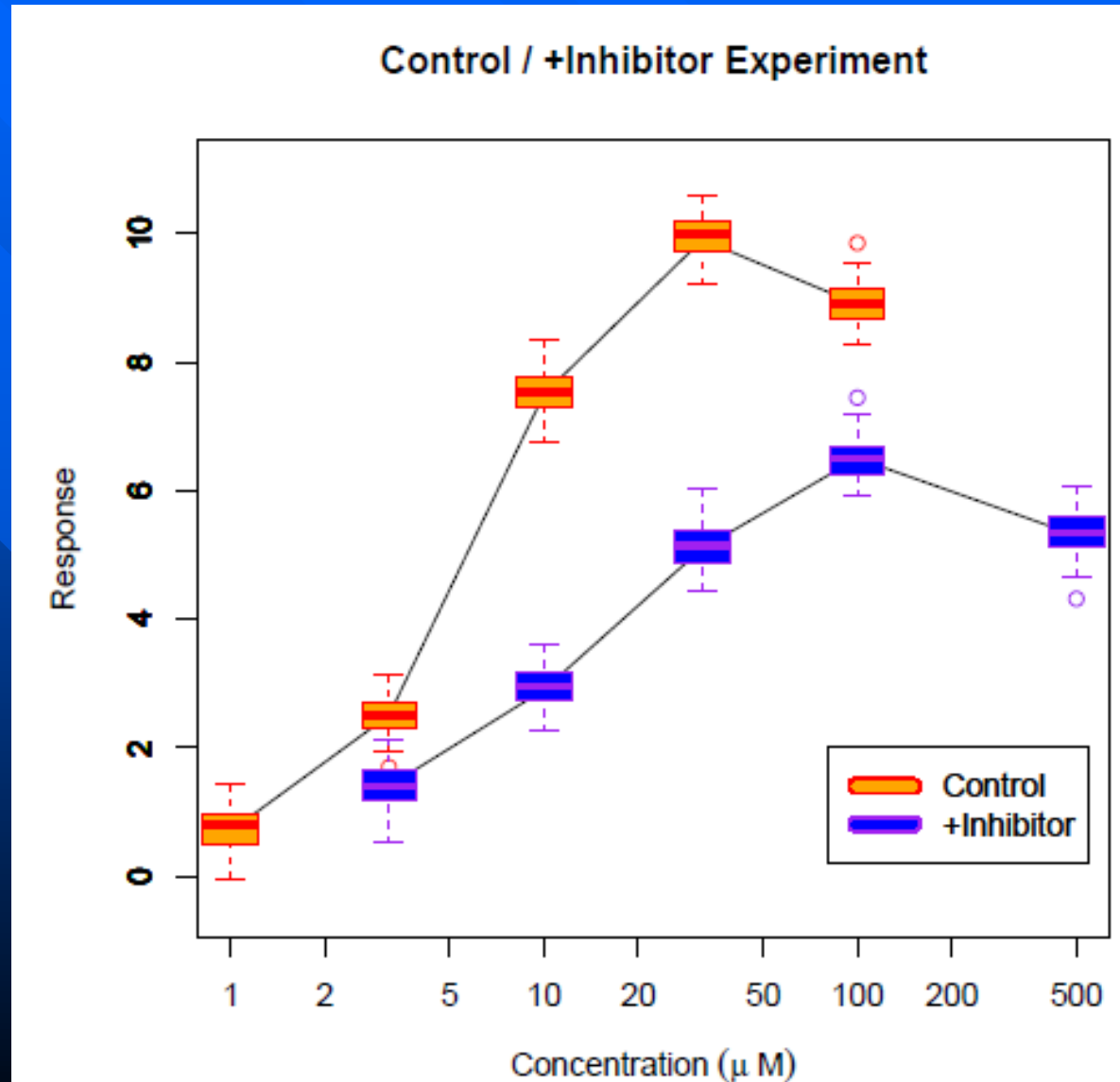
Rules violated:

- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.
- Rule 12: If it has been done well in the past, think of a new way to do it.

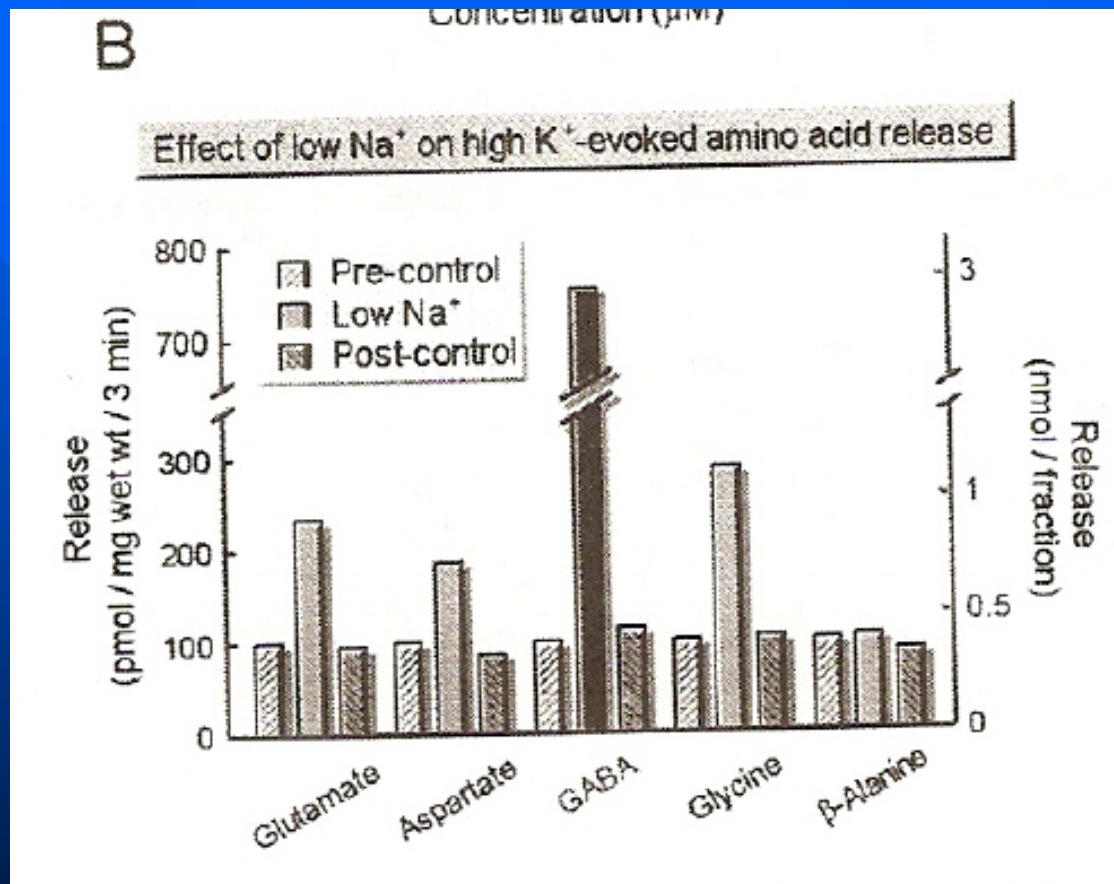
Computational Statistics 2002 (4)



Improved
Graph



Computational Statistics 2002 (5)

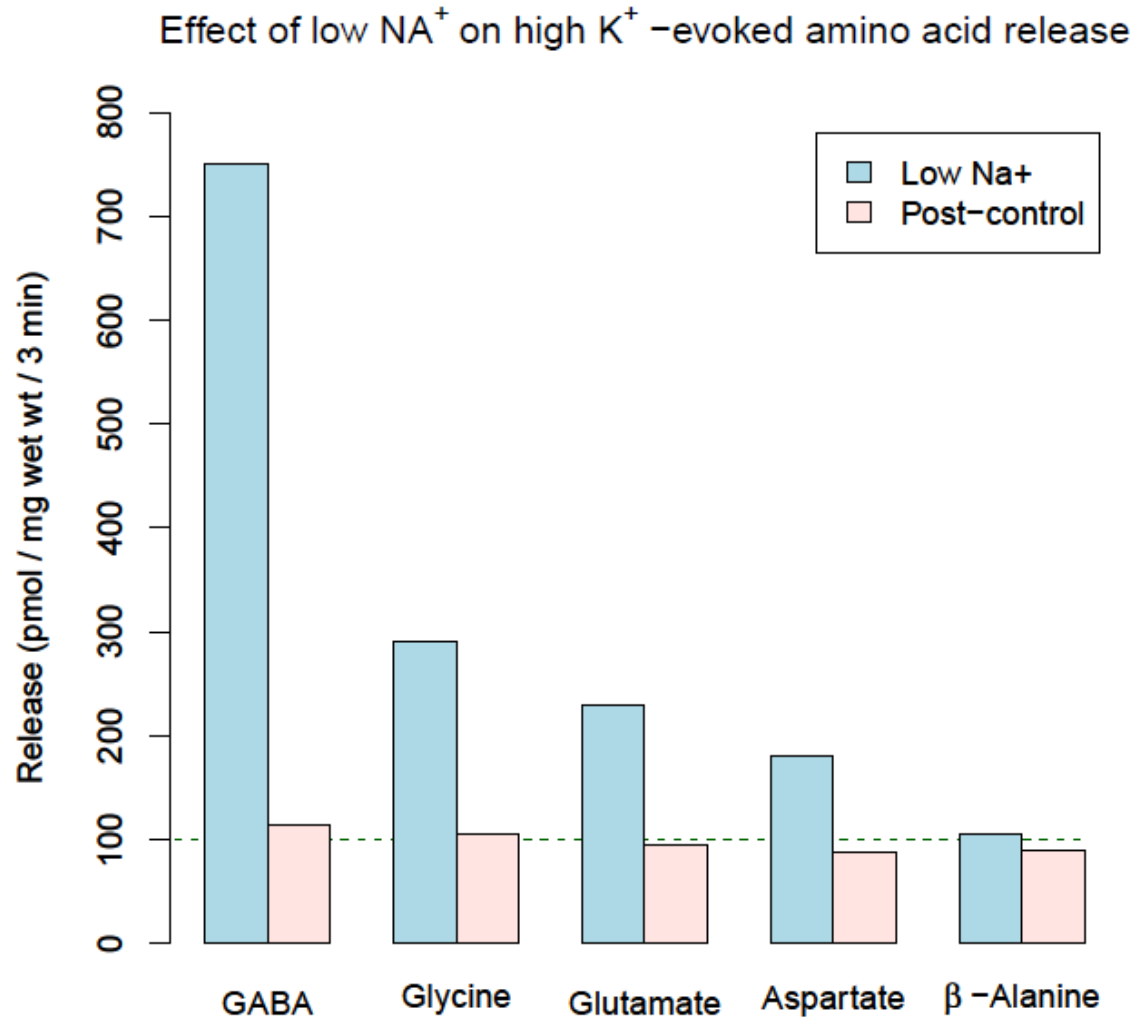
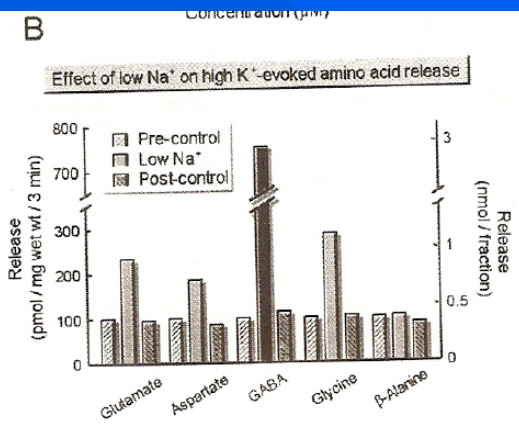


Computational Statistics 2002 (6)

Rules violated:

- Rule 6: Change scales in mid-axis.
- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.
- Rule 9: Alabama first!
- Rule 7: Emphasize the trivial (ignore the important).
- Rule 3: Ignore the visual metaphor altogether.

Computational Statistics 2002 (7)



Pre-control = 100

Improved
Graph

Computational Statistics 2016 (1)

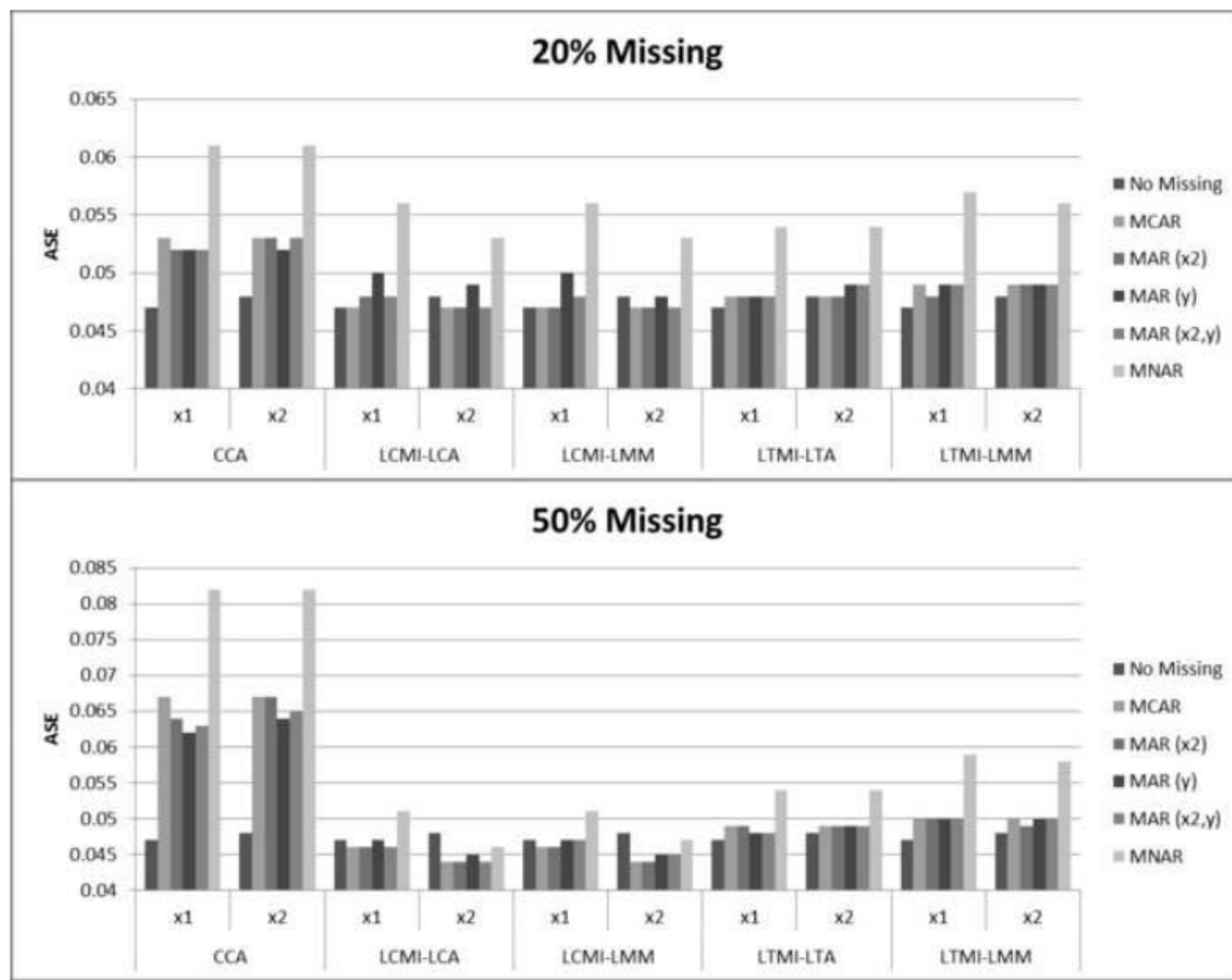


Figure 1. Comparison of asymptotic standard errors for all simulated models with count outcome Y . Results are stratified by percentage of missing data, method of dealing with missingness (CCA, LCMI, or LTMi methods), and type of missingness (MCAR, MAR, MNAR) for predictor X_1 including missingness and complete covariate X_2 .

Computational Statistics 2016 (2)

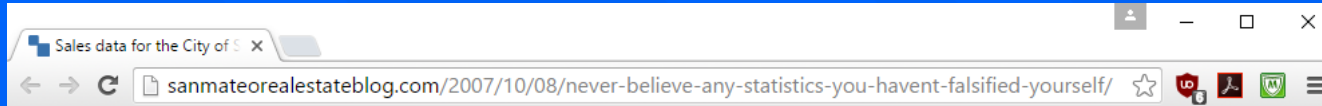
Rules violated:

- Rule 4: Only order matters.
- Rule 9: Alabama first!
- Rule 7: Emphasize the trivial (ignore the important).
- Rule 10: Label: (a) illegibly, (b) incompletely, (c) incorrectly, and (d) ambiguously.

Possible Explanations for Misleading Graphics

- Graphs should **intentionally** mislead!
- **Lack of Knowledge #1:**
Graphs constructed by non-statisticians, e.g.,
graphic/web designers.
- **Lack of Knowledge #2:**
Even teachers often do not know what the correct
graphs should look like.

Conclusion ???



Sales data for the City of San Carlos based on semi-annual data from 1998 to 2007

OCTOBER 8, 2007 BY VICKI MOORE

LEAVE A COMMENT



Having been in real estate 5 or so years, I never had any use for statistics. Put the house on the market, a week later it sold.

Now it's another story. Statistics are a crucial part of pricing homes for sale and for determining an appropriate offer price for a buyer. There's a constant discussion happening about what the market is doing and where it's doing it; statistics play a vital role in that discussion.

My client has this great quote on his website, **"Never believe any statistics you haven't falsified yourself."** Just so you know, I didn't do anything to these statistics to manipulate them. It's a straight pull of information from the Multiple Listing Service for all neighborhoods within the City of San Carlos based on semi-annual data from 1998 to 2007.

Current Work

- Interact with the *American Statistical Association (ASA) / National Council of Teachers of Mathematics (NCTM) Joint Committee*.
- Revisit winning posters from the annual Poster Competition and Project Competition (<http://www.amstat.org/education/posterprojects>).
- Point out common mistakes in graphs from winning posters
- Revisit guidelines and provide educational materials (webinar, online booklet for students, teachers, and judges).

Sources for Constructing Better Graphics: Books & Articles

- Robbins, N. B. (2013). *Creating More Effective Graphs*. Chart House: Ramsey, NJ. (reprinted from Wiley 2005)
- Su, Y-S. (2008). “It’s easy to produce chartjunk using Microsoft Excel 2007, but hard to make good graphs.” *Computational Statistics and Data Analysis* 52:4594-4601.
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information*, 2nd edition. Graphics Press: Cheshire, CT. (first edition 1983)
- Wainer, H. (2000). *Visual Revelations: Graphical Tales of Fate and Deception from Napoleon Bonaparte to Ross Perot*. Psychology Press: London, UK. (reprinted from Copernicus 1997)

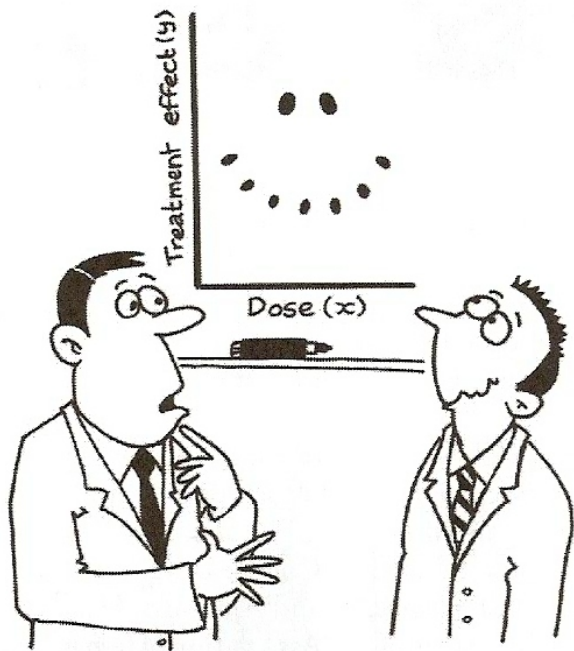
Sources for Constructing Better Graphics:

Blogs

- Camoes, J. “ExcelCharts,” <http://www.excelcharts.com/blog/>
- Peltier, J. “Peltier Tech Blog— Peltier Tech Excel Charts and Programming Blog,” <http://peltiertech.com/WordPress/>
- Robbins, N. B. “Effective Graphs,” <http://www.forbes.com/sites/naomirobbins/>
- Robbins, N. B. (February 28, 2012). “Misleading Graphs: Displaying a Change in One Variable Using Area or Volume,” Forbes.com LLC. <http://www.forbes.com/sites/naomirobbins/2012/02/28/misleading-graphs-displaying-a-change-in-one-variable-using-area-or-volume/>
- Wyatt, S. (March 31, 2014). “Dishonest Fox Charts: Obamacare Enrollment Edition,” Media Matters for America. <http://mediamatters.org/blog/2014/03/31/dishonest-fox-charts-obamacare-enrollment-editi/198679>

Sources for Constructing Better Graphics: For Teachers

- Heiberger, R. M., Robbins, N. B., Symanzik, J. (2014). “Statistical Graphics Recommendations for the ASA/NCTM Annual Poster Competition and Project Competition”, JSM Proceedings, American Statistical Association, Alexandria, VA.
- Symanzik, J., Robbins, N. B., Heiberger, R. M. (2014). “Observations from the winners of the 2013 Statistics Poster Competition — Praise and future improvements.” The Statistics Teacher Network, 83:2–5.
<http://www.amstat.org/education/stn/pdfs/STN83.pdf>



"It's a non-linear pattern with outliers.....but for some reason I'm very happy with the data."

A CAUSE-commissioned cartoon that is part of the CAUSEweb collection and available for free noncommercial use by statistics teachers. Cartoon by John Landers ©. Provided by permission.

Questions ???

— or —

send e-mail to:

symanzik@math.usu.edu

From: Amstat News, January 2009, p. 25