DESIGN GUIDELINE

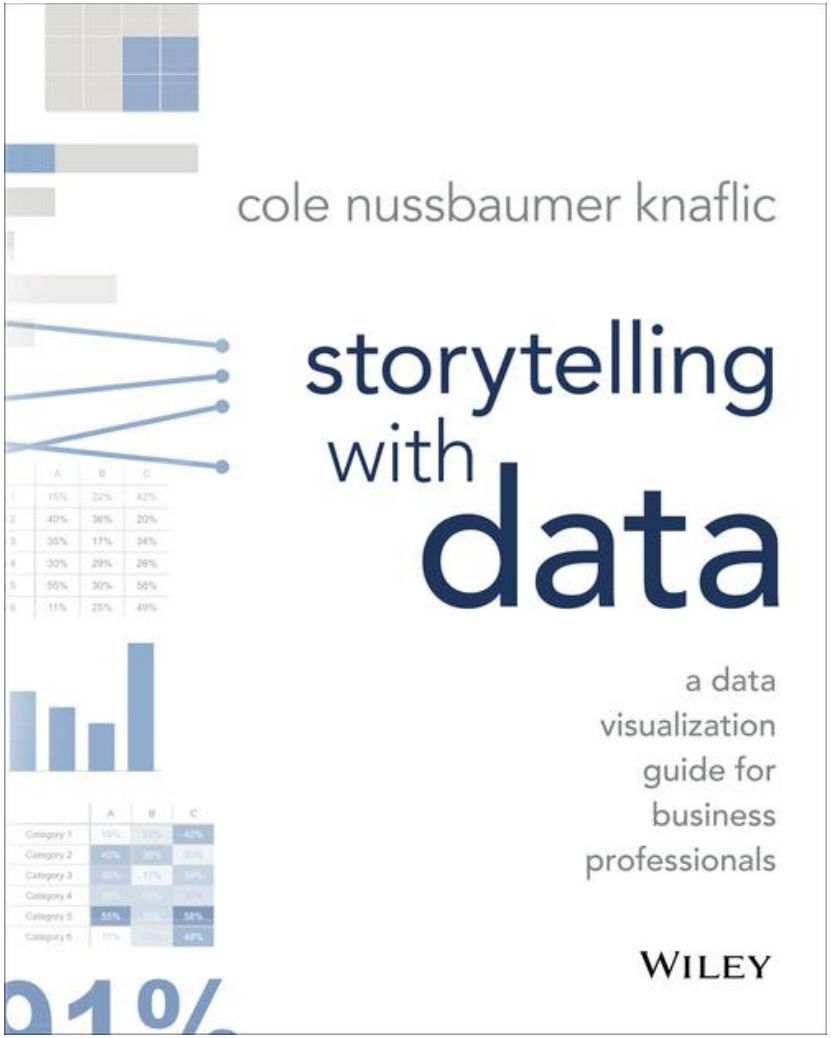
PRINCIPLES

#1: The visualization should show all of the data and only the data

PRINCIPLES

#2: Use the best visual channel available for the most important aspect of the data

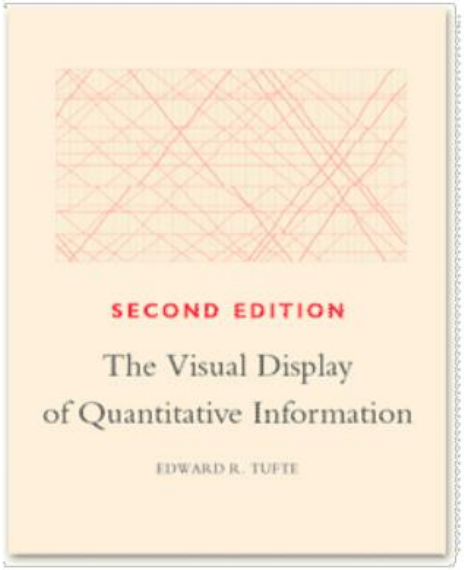
MATERIALS

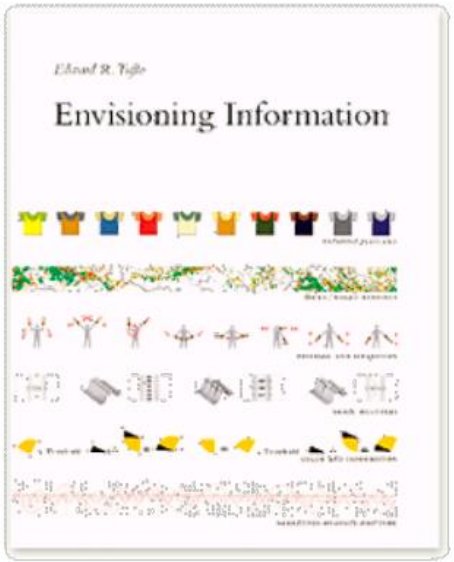

cole nussbaumer knaflic
storytelling with data
a data visualization guide for business professionals
WILEY

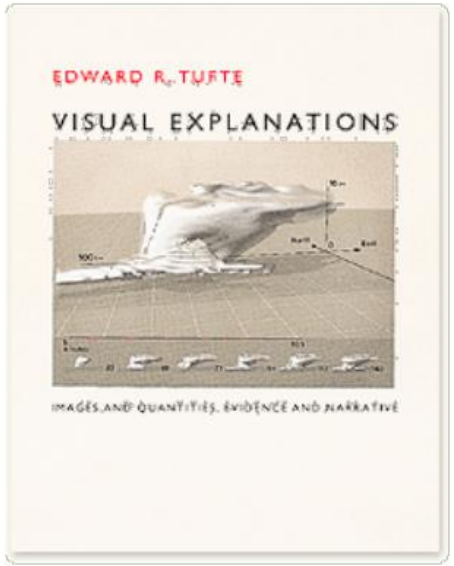
MATERIALS

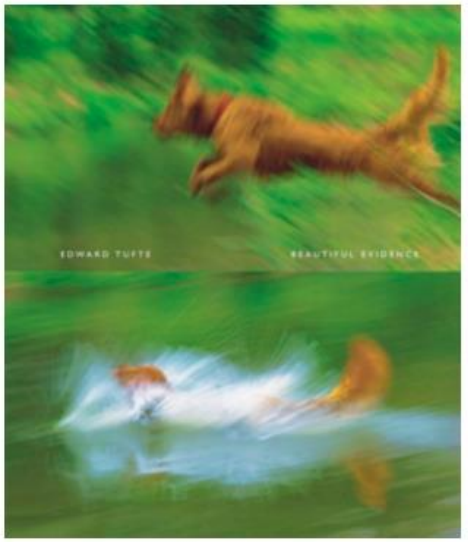
Edward Tufte

Author of fundamental books on visualization









https://en.wikipedia.org/wiki/Edward_Tufte

THE VISUAL DISPLAY OF QUANTITATIVE INFORMATION

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

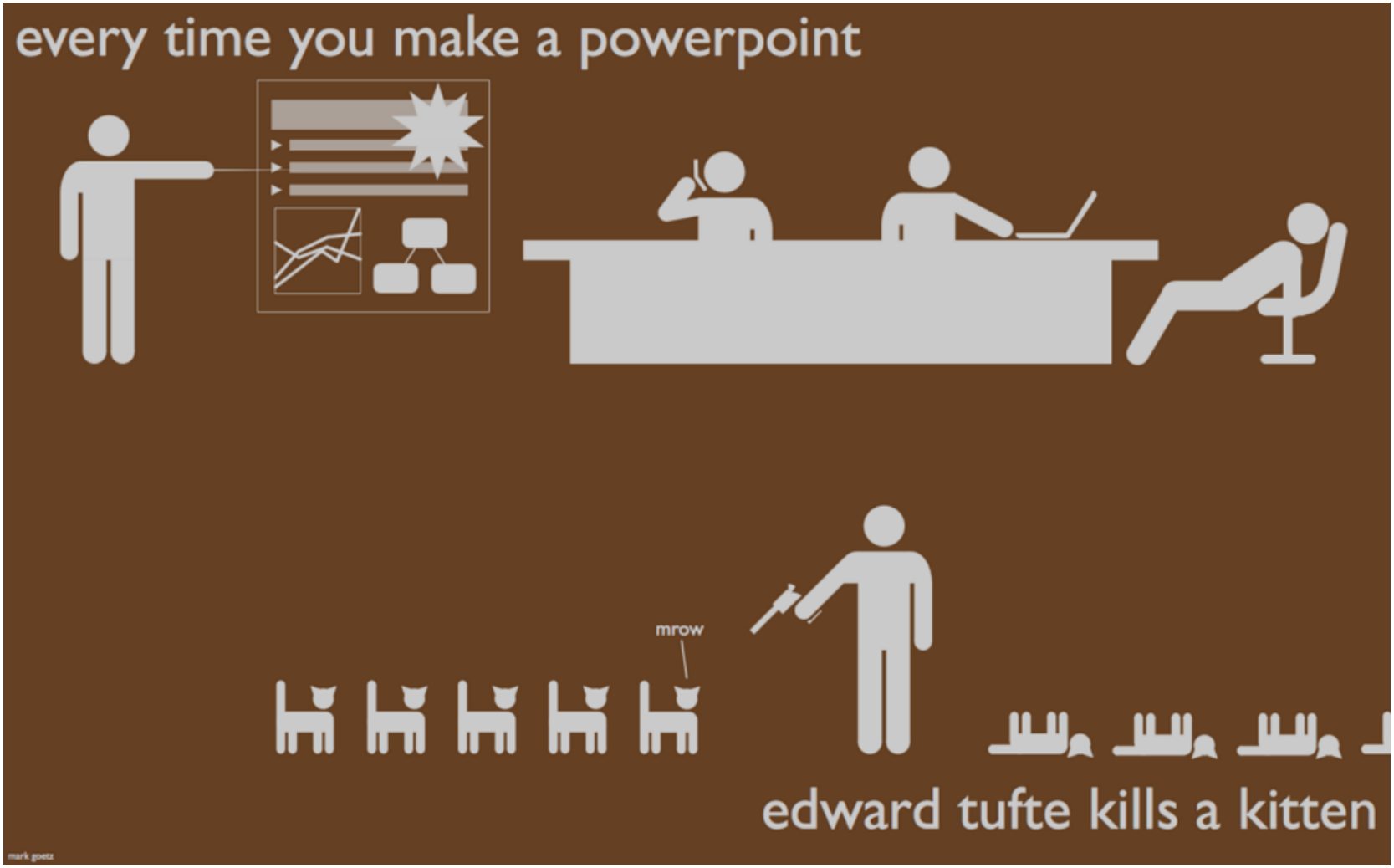
Edward Tufte

MORE GUIDELINES

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

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

And many more in **The Visual Display of Quantitative Information**


every time you make a powerpoint
mrow
edward tufte kills a kitten
mark goetz

TUFTE'S INTEGRITY PRINCIPLES

Show **data variation,** not design variation

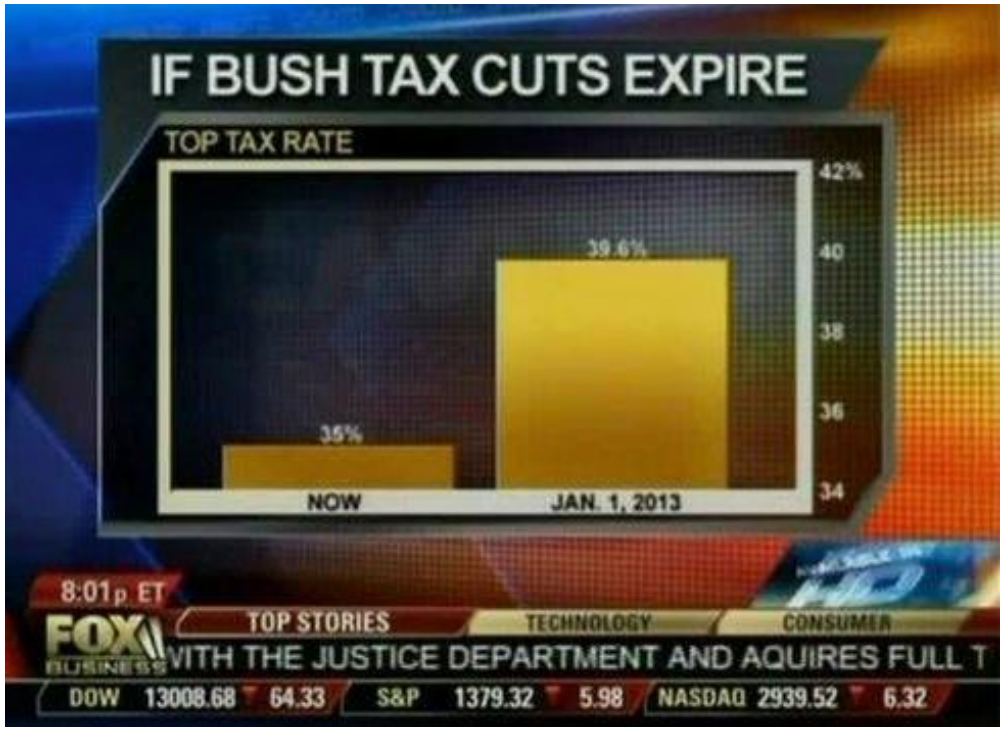
Clear, detailed, and thorough **labeling** and **appropriate scales**

Size of the **graphic effect** should be **directly proportional to the numerical quantities** ("lie factor")

THE LIE FACTOR

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

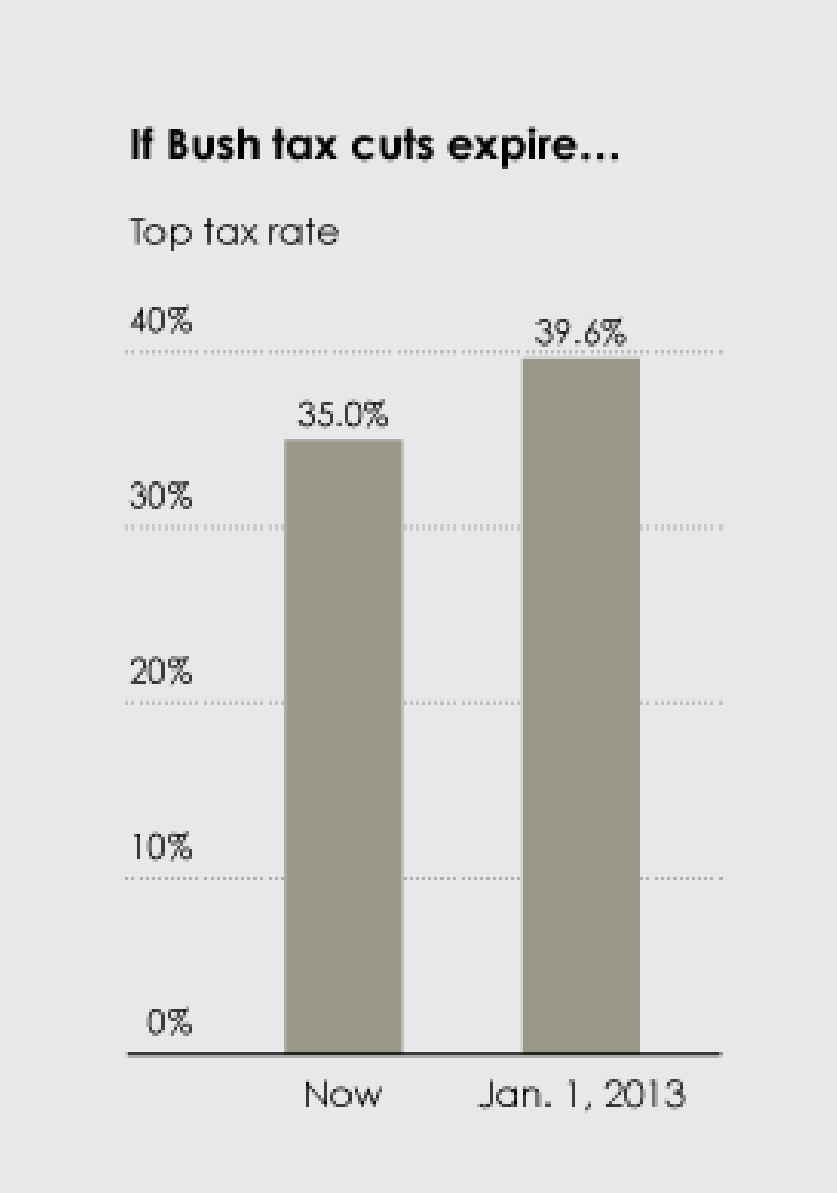
LIE FACTOR - GRAPHICAL INTEGRITY



Magnitude in data must correspond to magnitude of mark

Effect in Data: factor 1.14
Effect in Graphic: factor 5
Lie Factor: 5/1.14 = 4.38

SCALE DISTORTIONS



WHAT'S WRONG?

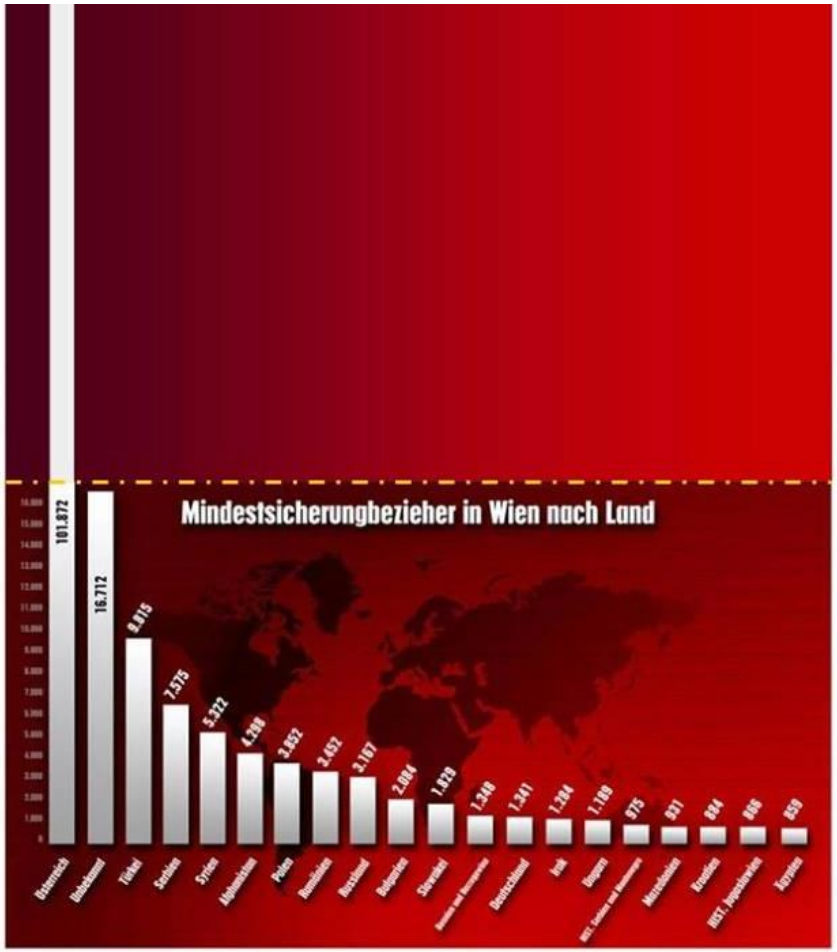
Viele Bezieher mit "ungeklärter Staatsbürgerschaft"

Die größte Gruppe in der Liste der Mindestsicherungsbezieher ist aber jene der "ungeklärten Staatsbürgerschaft". Dass es sich bei den 16.712 Personen um

WHAT'S WRONG?

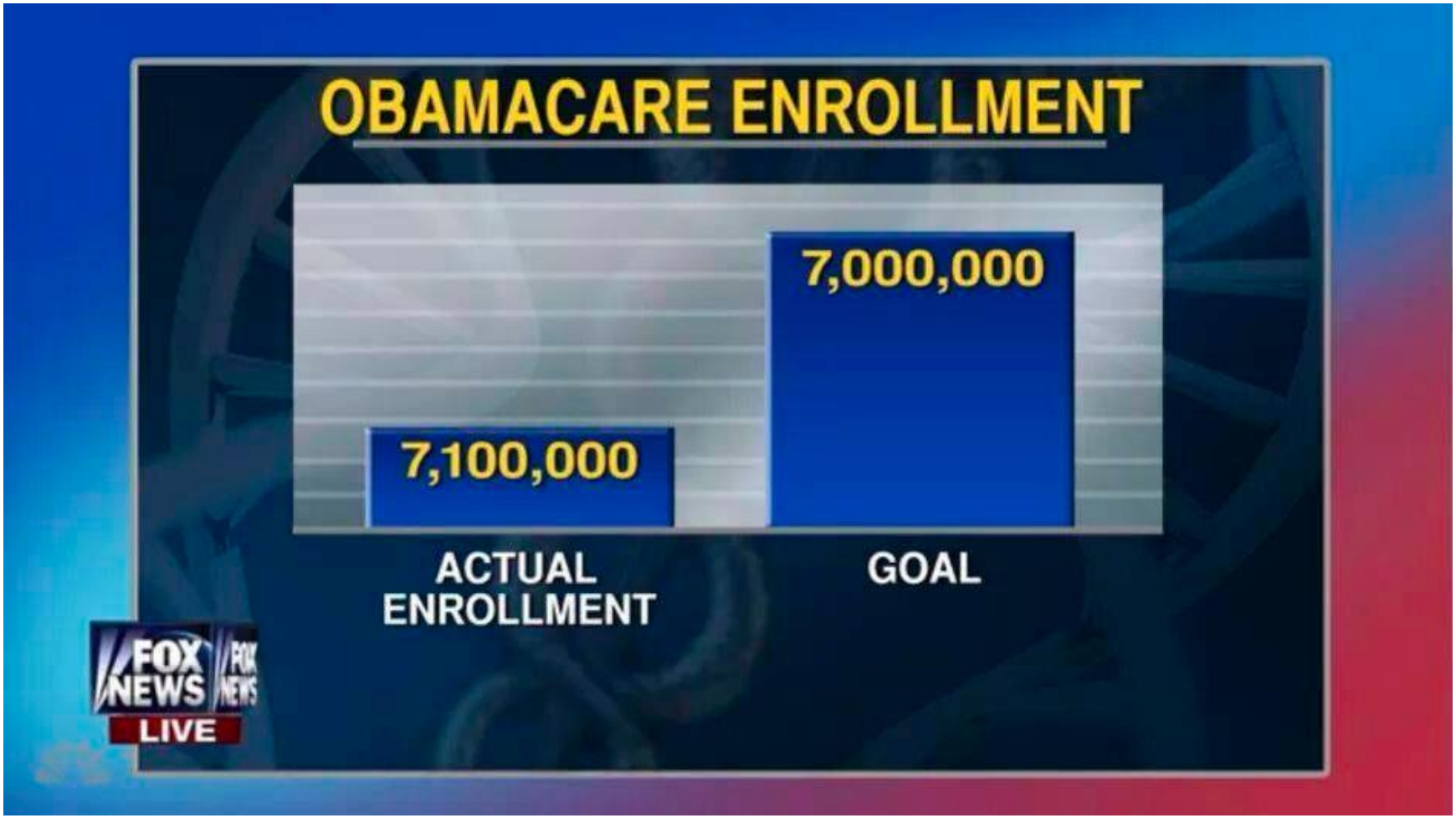
Viele Bezieher mit "ungeklärter Staatsbürgerschaft"

Die größte Gruppe in der Liste der Mindestsicherungsbezieher ist aber jene der "ungeklärten Staatsbürgerschaft". Dass es sich bei den 16.712 Personen um

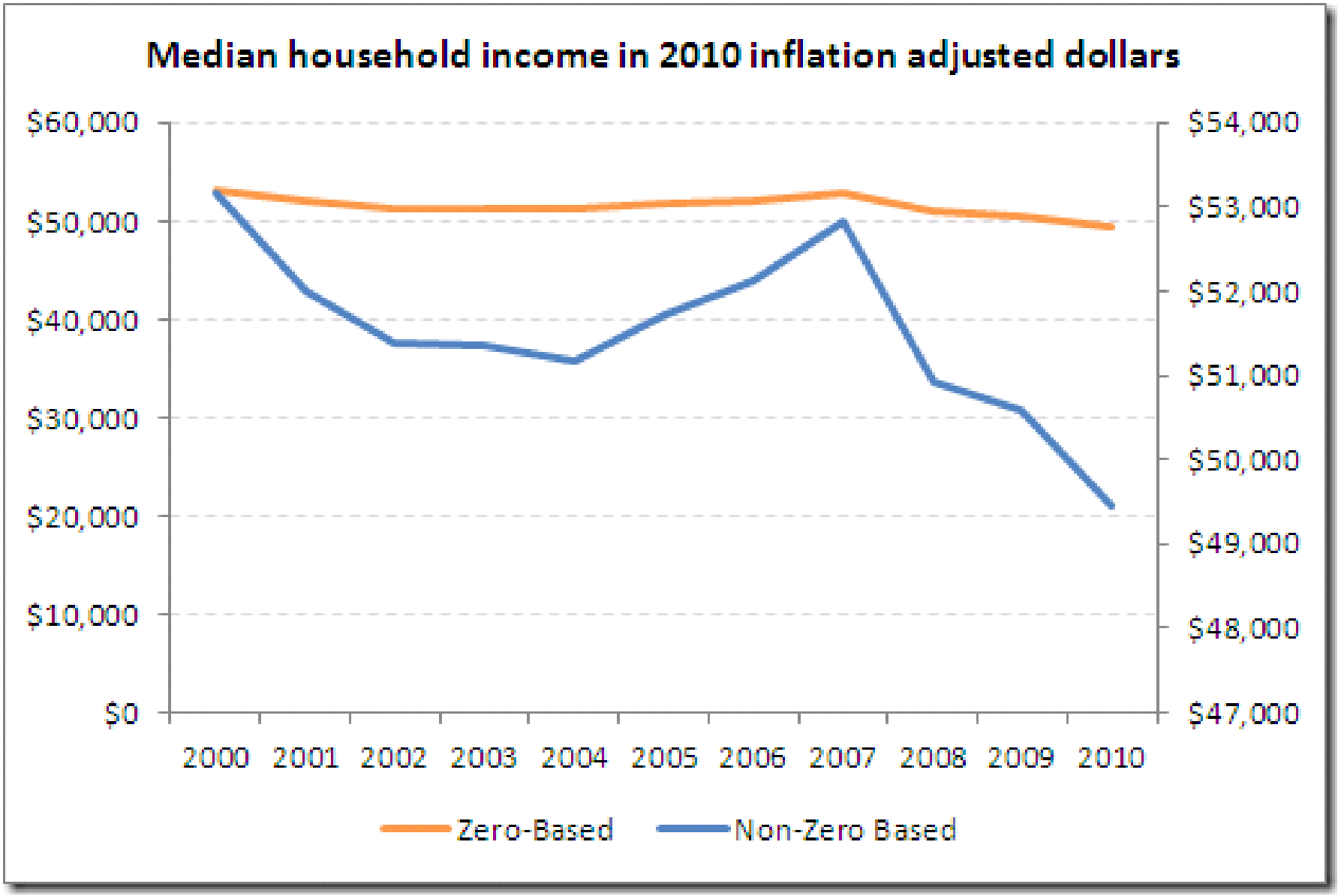


Viele Bezieher mit "ungeklärter Staatsbürgerschaft"

Die größte Gruppe in der Liste der Mindestsicherungsbezieher ist aber jene der "ungeklärten Staatsbürgerschaft". Dass es sich bei den 16.712 Personen um


OBAMACARE ENROLLMENT
7,000,000
7,100,000
ACTUAL ENROLLMENT
GOAL
FOX NEWS
FOX NEWS
LIVE

START AT 0?


Median household income in 2010 inflation adjusted dollars
$60,000
$50,000
$40,000
$30,000
$20,000
$10,000
$0
$54,000
$53,000
$52,000
$51,000
$50,000
$49,000
$48,000
$47,000
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Zero-Based
Non-Zero Based

Use a baseline that shows the data, not the zero-point

Think about: what is a meaningful baseline?

SCALE AT 0

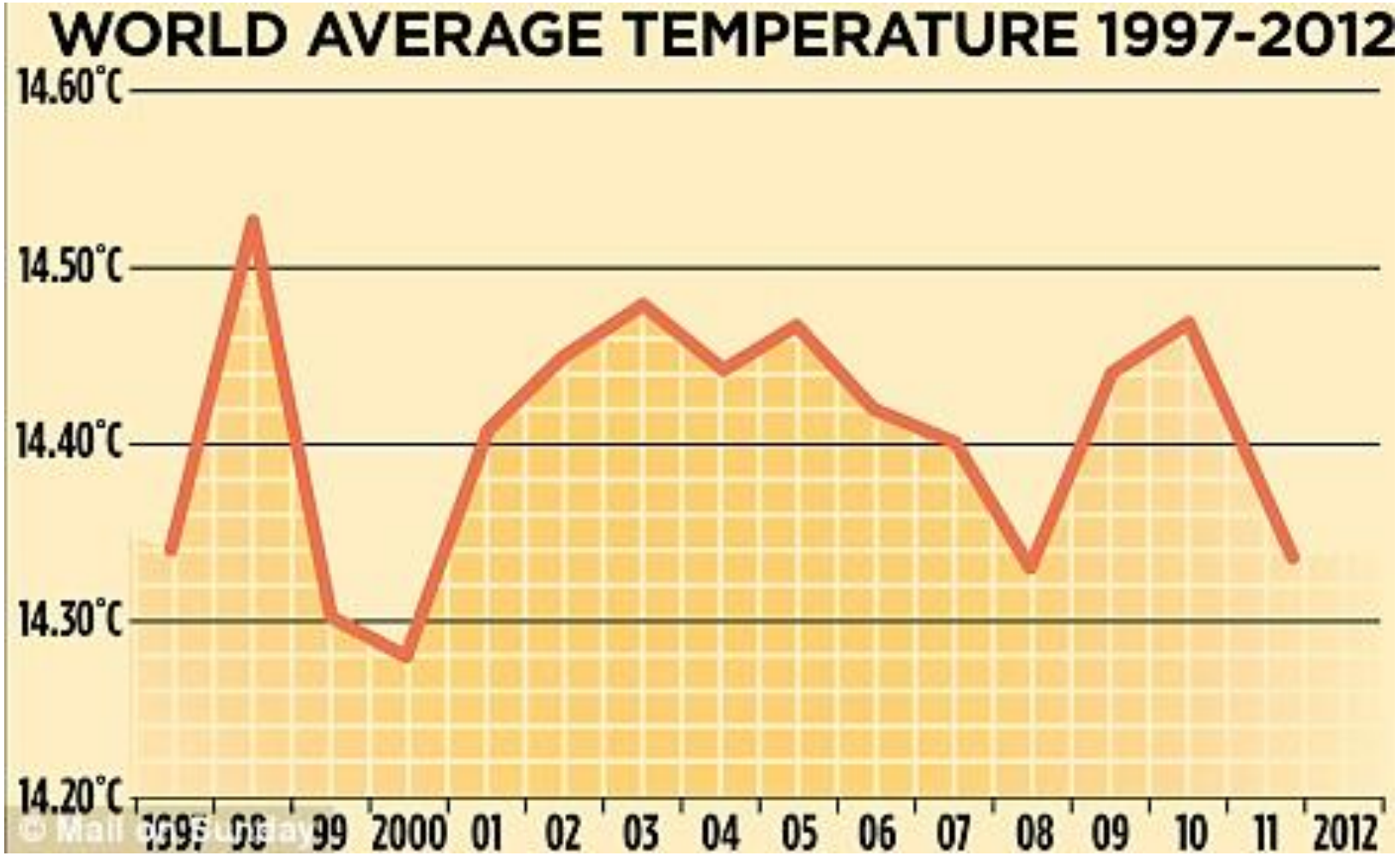
FRAMING

Vis can be used to lie

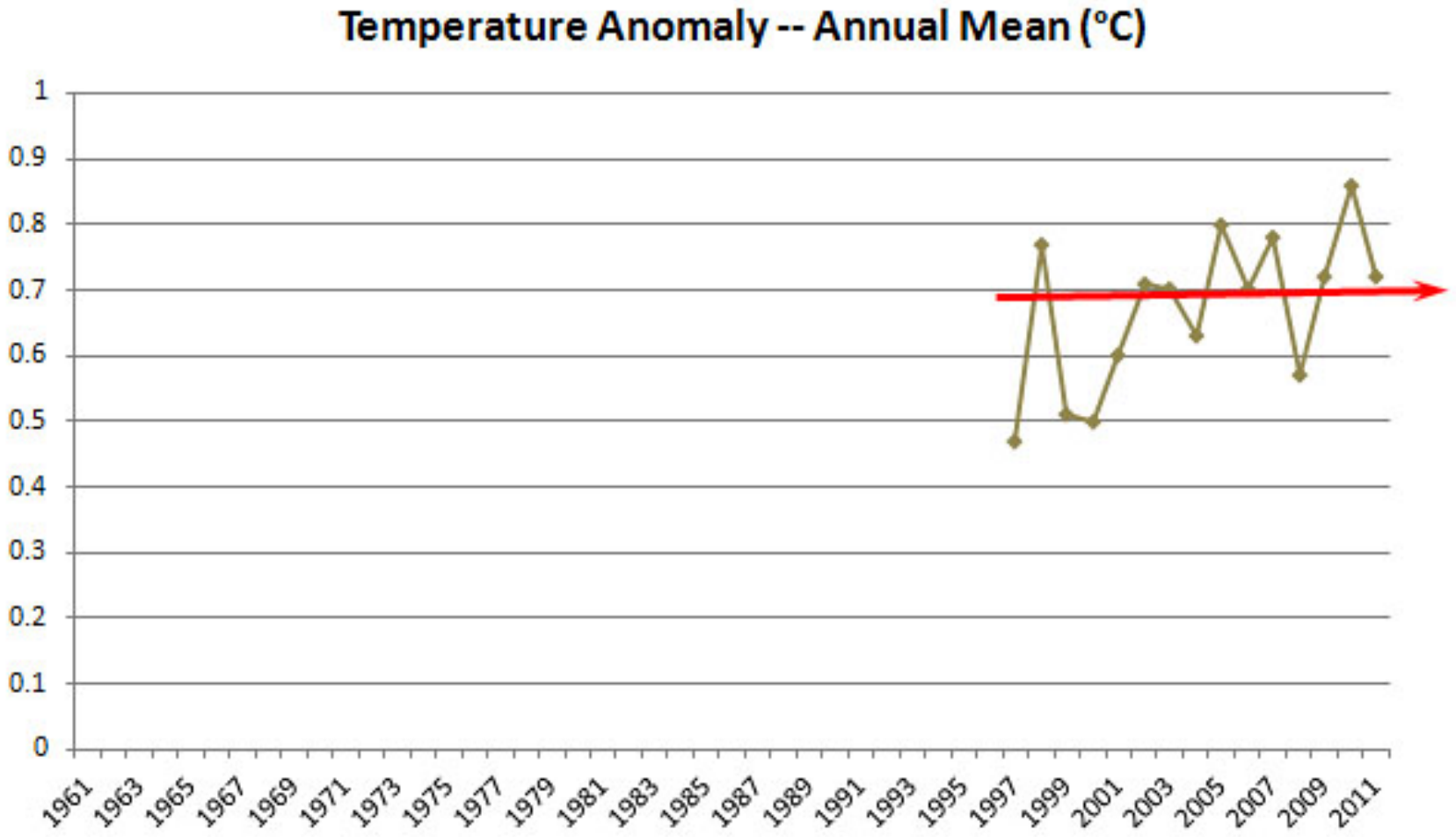
as language or statistics

When showing something, make sure that you're faithful to the data

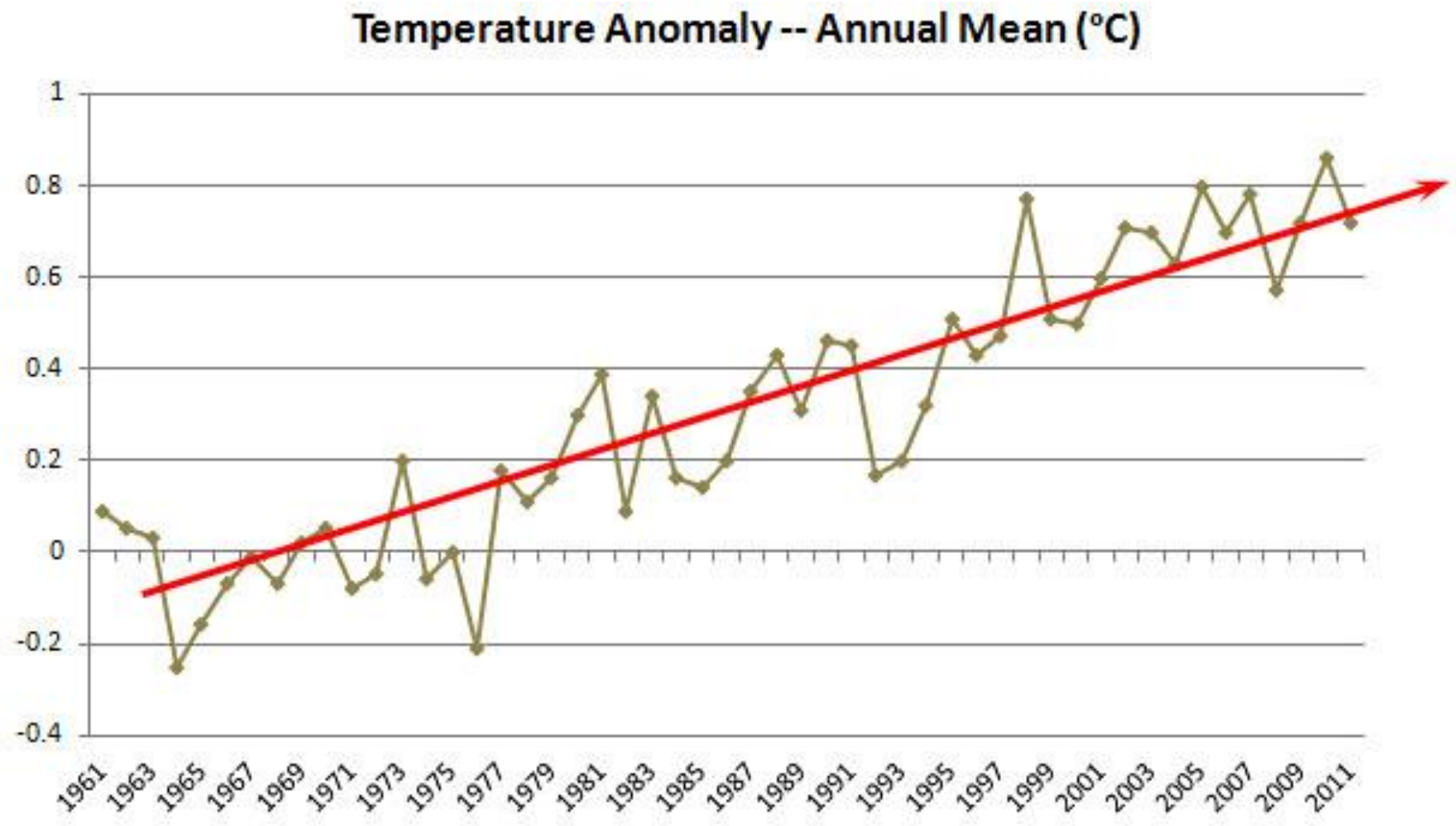
GLOBAL WARMING?



The Daily Mail, UK, Jan 2012


Temperature Anomaly -- Annual Mean (°C)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011

GLOBAL WARNING – FRAME THE DATA



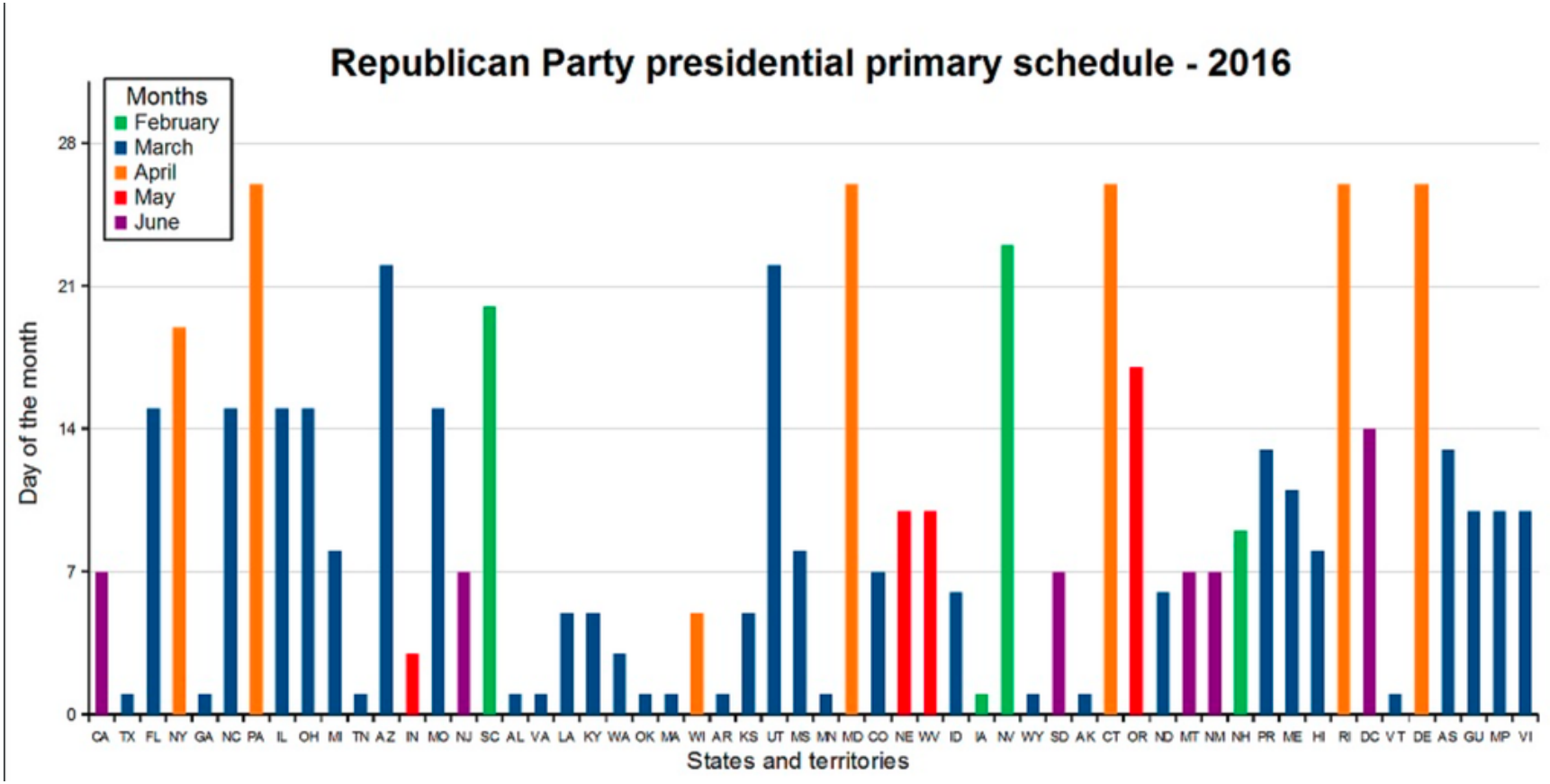
WHAT'S WRONG


How 2012 Stacks Up
The Warmest Years on Record
Contiguous U.S.
55.3
54.31
54.32
54.14
53.93
53.81
55.5°
55.0°
54.5°
54.0°
53.5°
53.0°
1921
1999
1934
2006
1998
2012
Source: NOAA's National Climatic Data Center - State of the Climate National Overview
CLIMATE CENTRAL

SCALE DISTORTIONS IN TEMPORAL DATA

SCALE DISTORTIONS IN TEMPORAL DATA

WHAT'S WRONG?

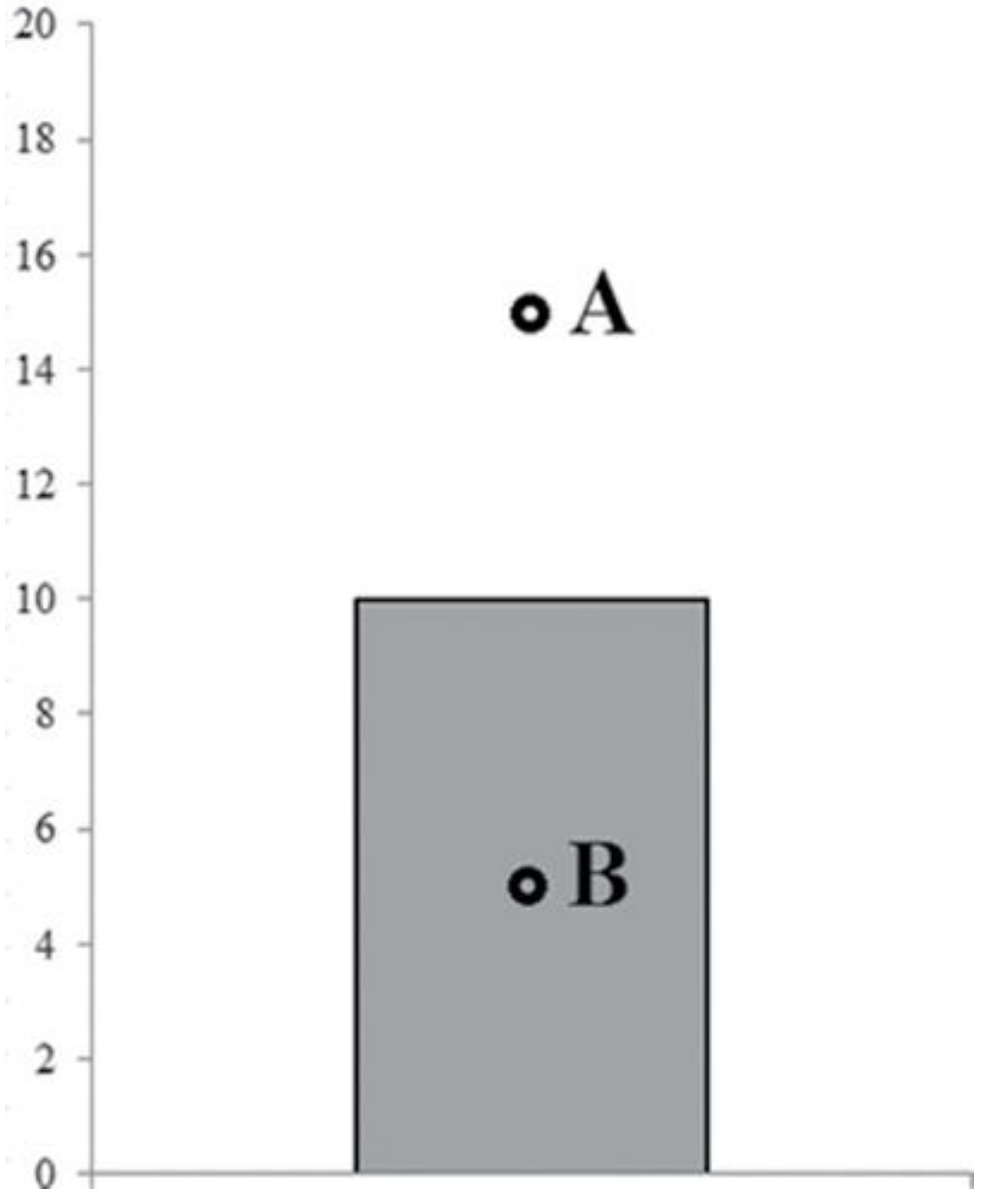

Republican Party presidential primary schedule - 2016
Months
February
March
April
May
June
Day of the month
0
7
14
21
28
CA TX FL NY GA NC PA IL OH MI TN AZ IN MO NJ SC AL VA LA KY WA OK MA WI AR KS UT MS MN MD CO NE WV ID IA NV WY SD AK CT OR ND MT NM NH PR ME HI RI DC VT DE AS GU MP VI
States and territories

BIASES

Height of the Bar encodes mean of a distribution

Which value is more likely to belong to the distribution?

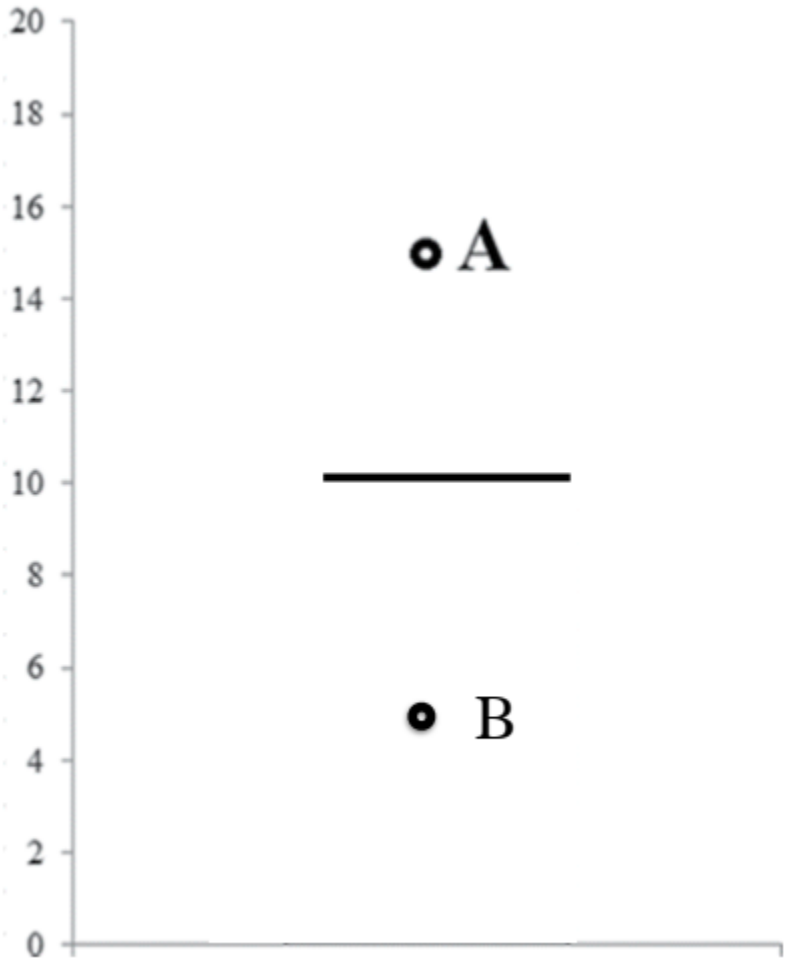
A or B?



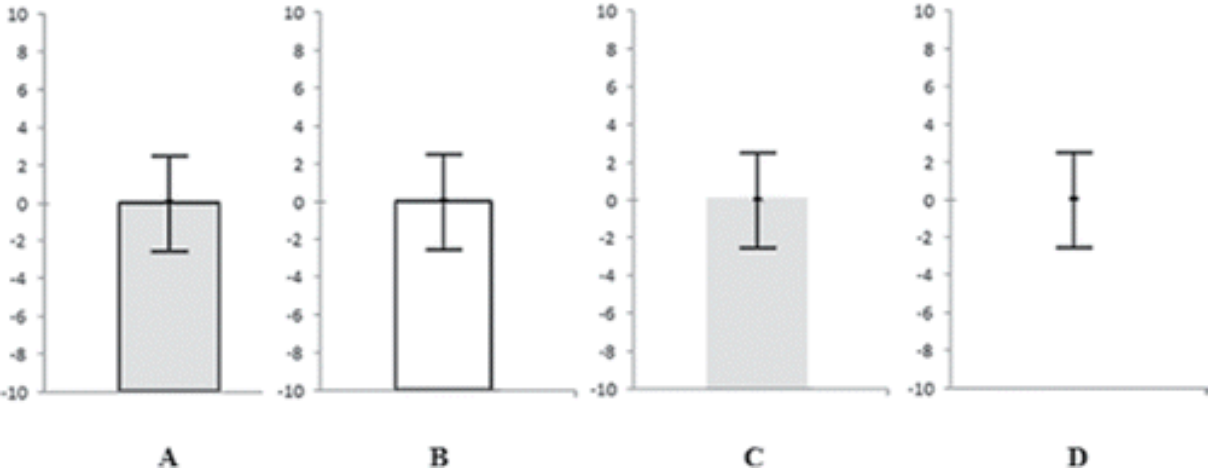
BIASES

We can plot the data faithfully, but still perceive it wrongly!

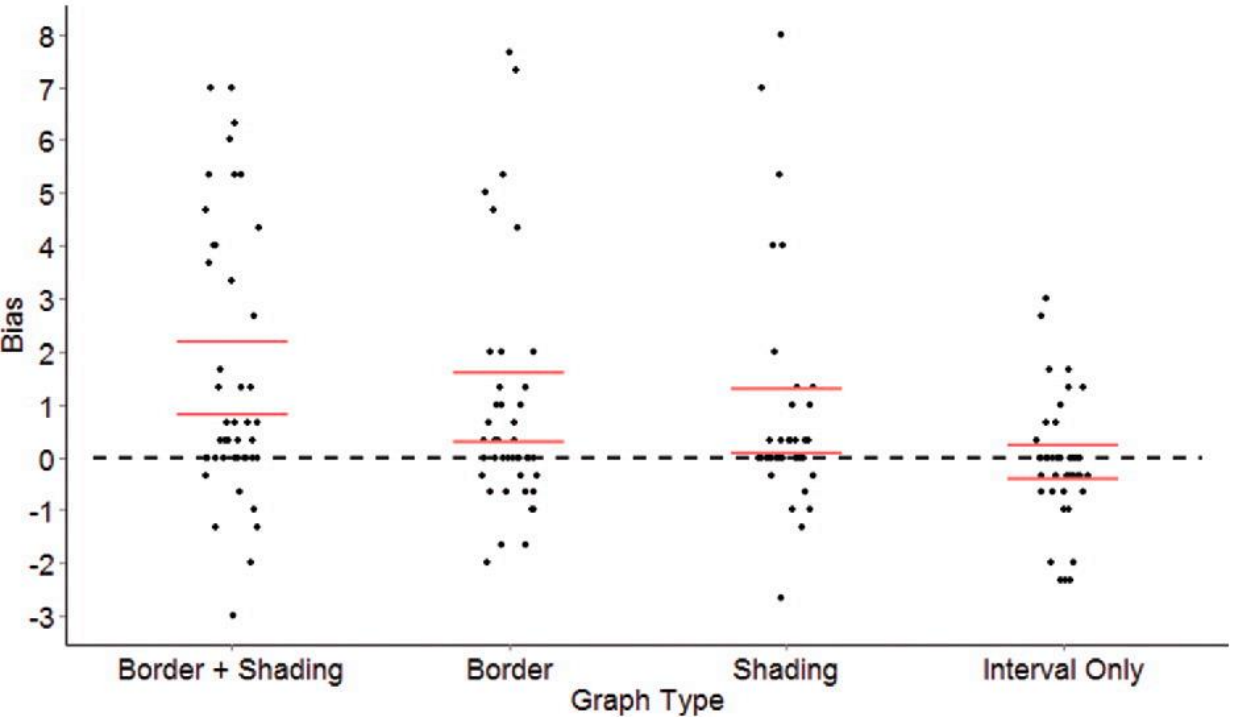
Mean of a distribution with a bar



WITHIN THE BAR BIAS



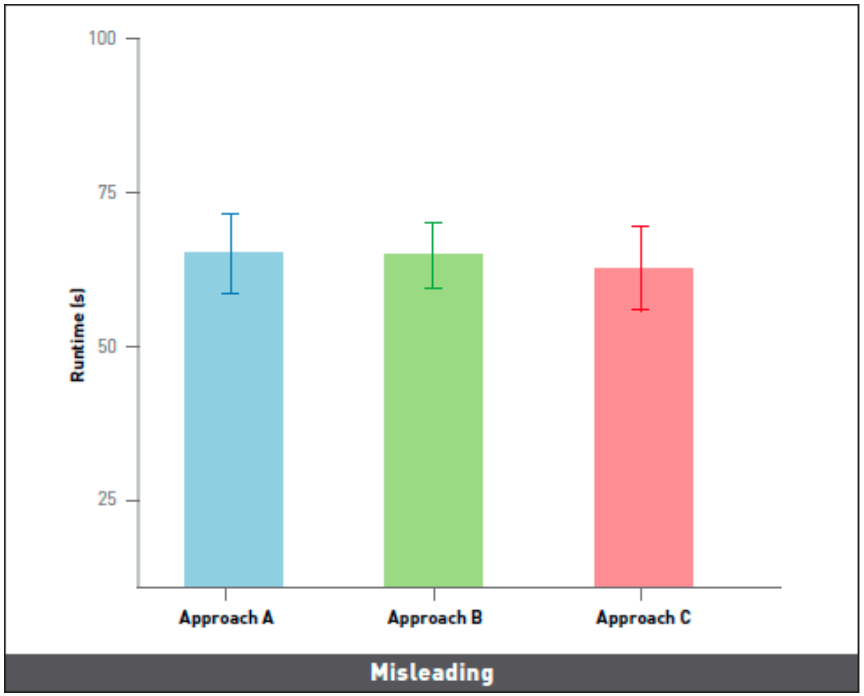
Experimental Conditions

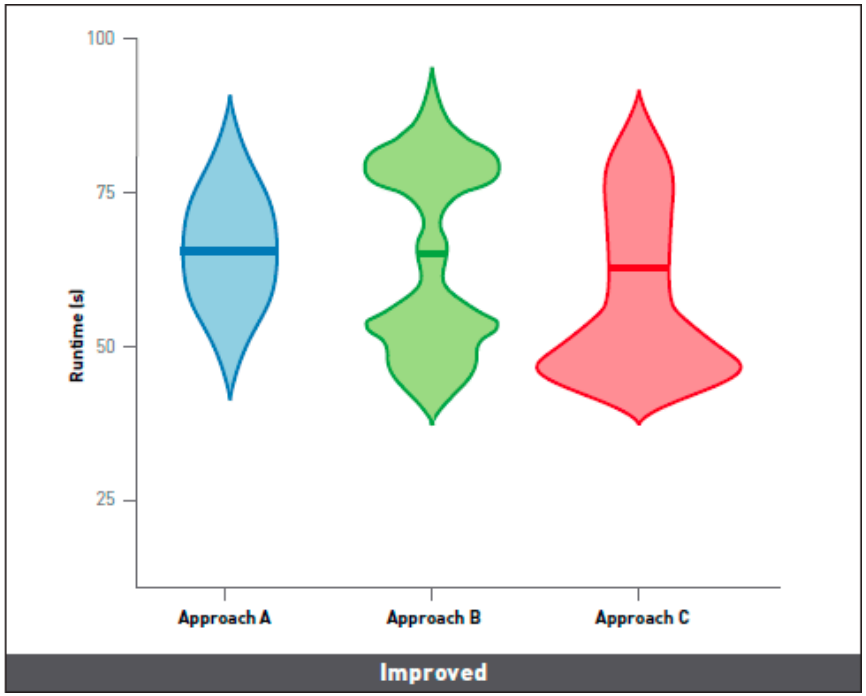


Results

Christopher S. Pentoney & Dale E. Berger (2016) Confidence Intervals and the Within-the-Bar Bias, The American Statistician, 70:2, 215-22

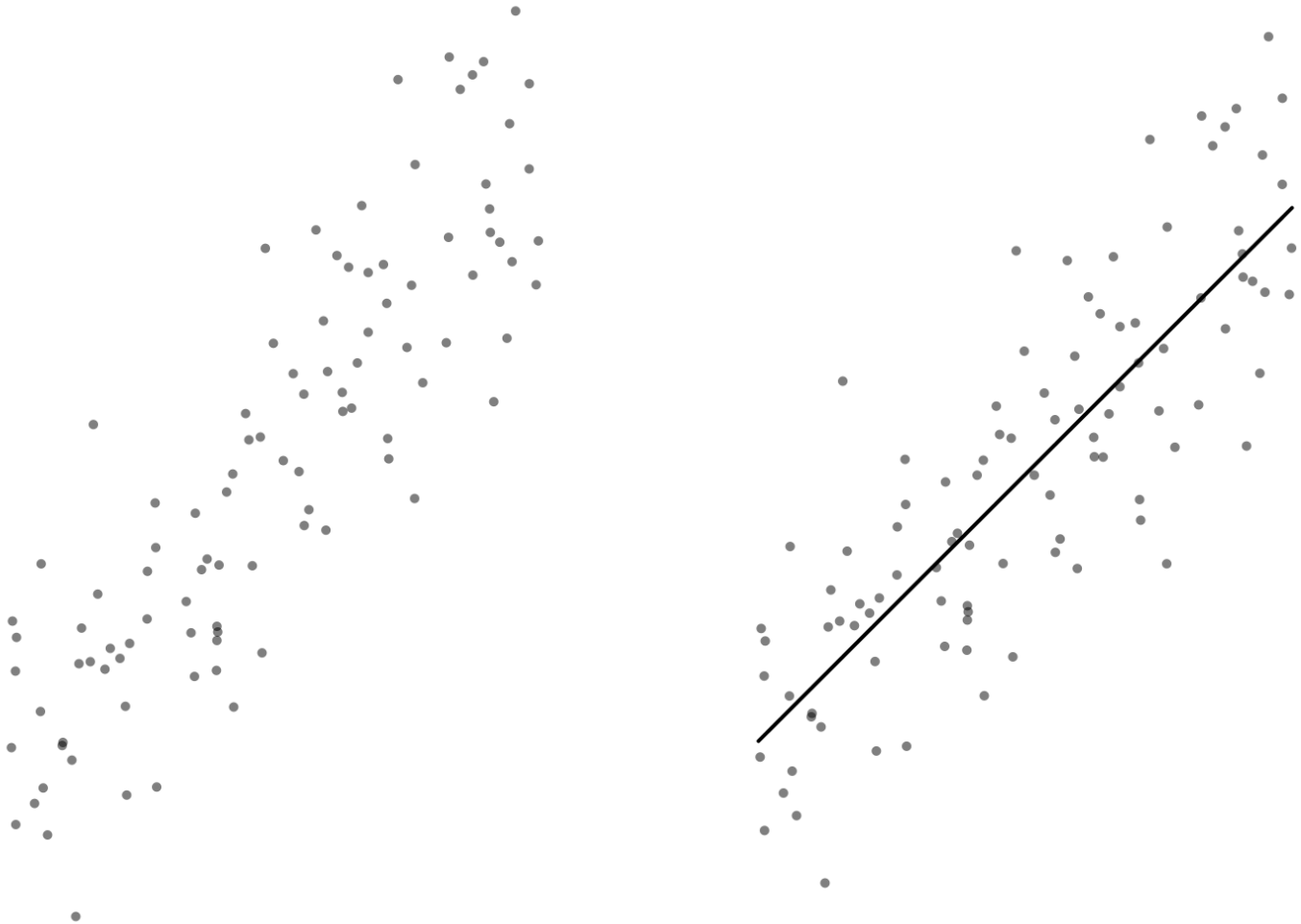
CAREFUL WHEN DESIGNING AGGREGATED CHARTS



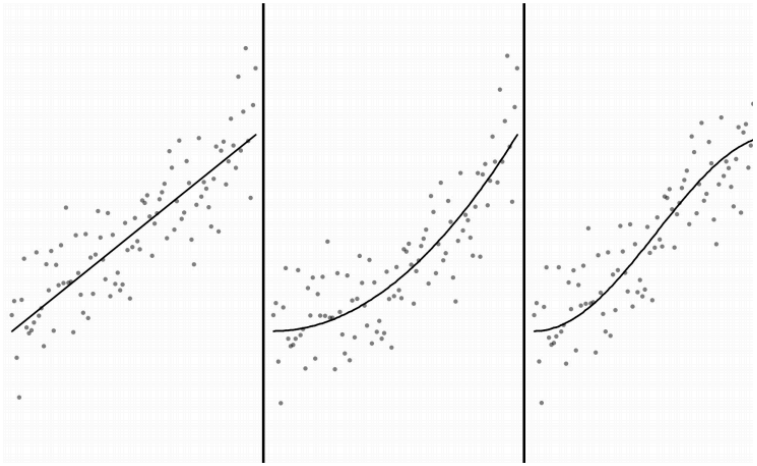


Should use median and confidence interval only when the data is normal distribution

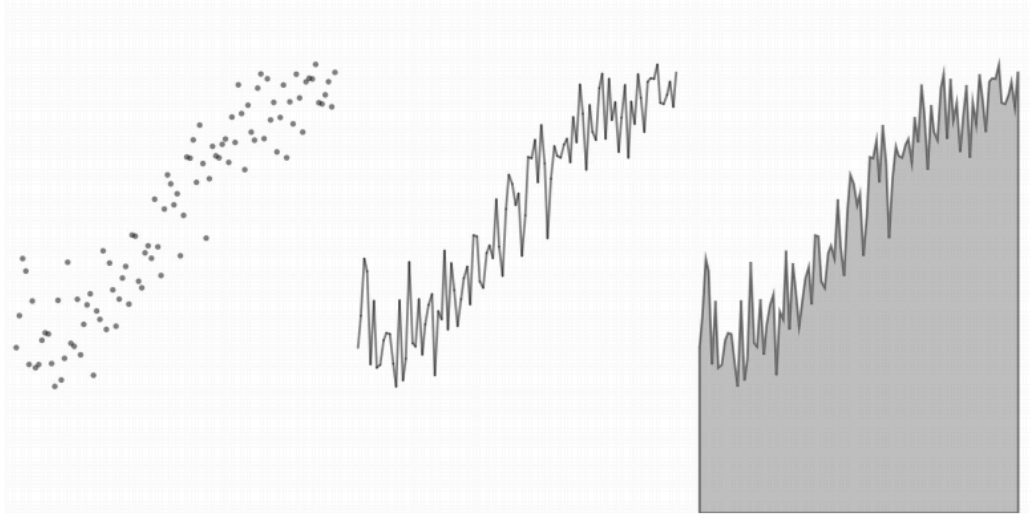
WHAT'S THE TRENDLINE



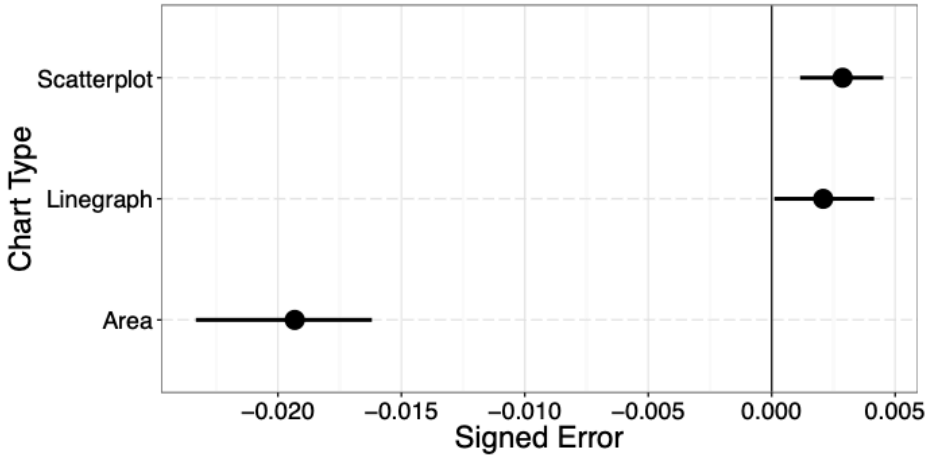
REGRESSION BY EYE



We’re good at spotting trends

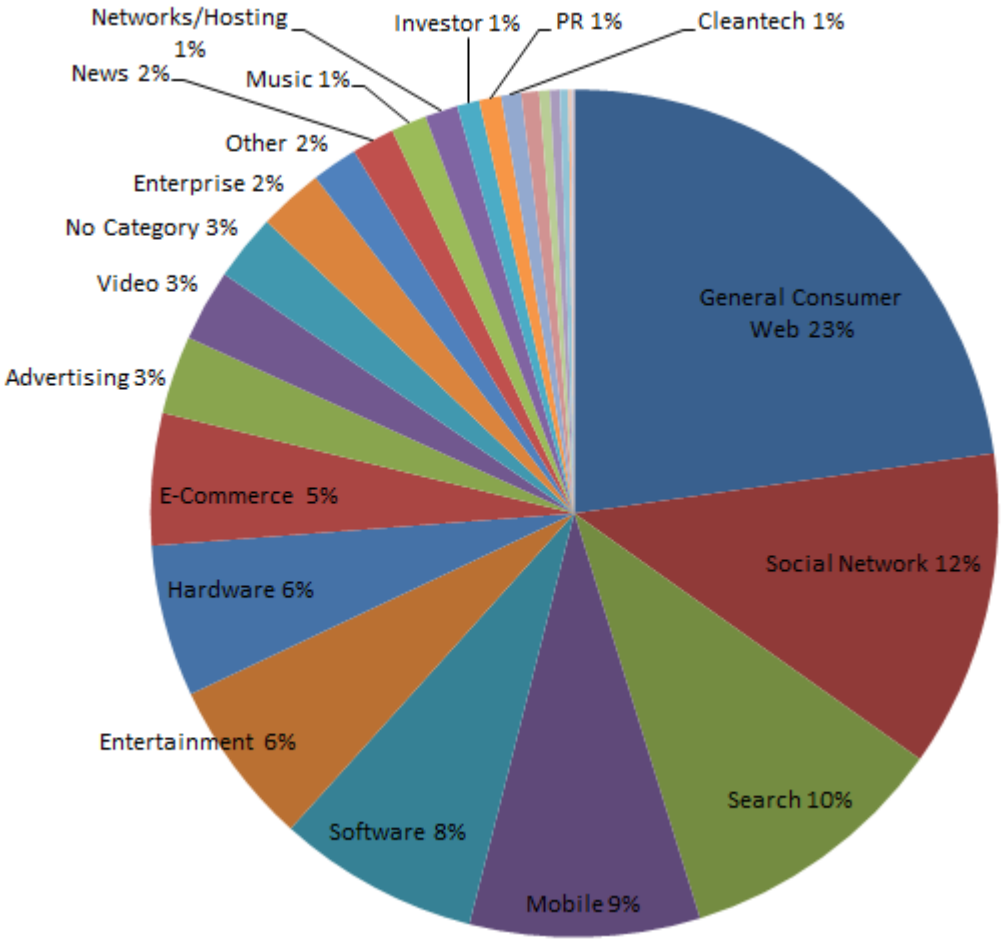


But the wrong vis technique can deceive us



http://idl.cs.washington.edu/files/2017-RegressionByEye-CHI.pdf

DEATH TO PIE CHARTS

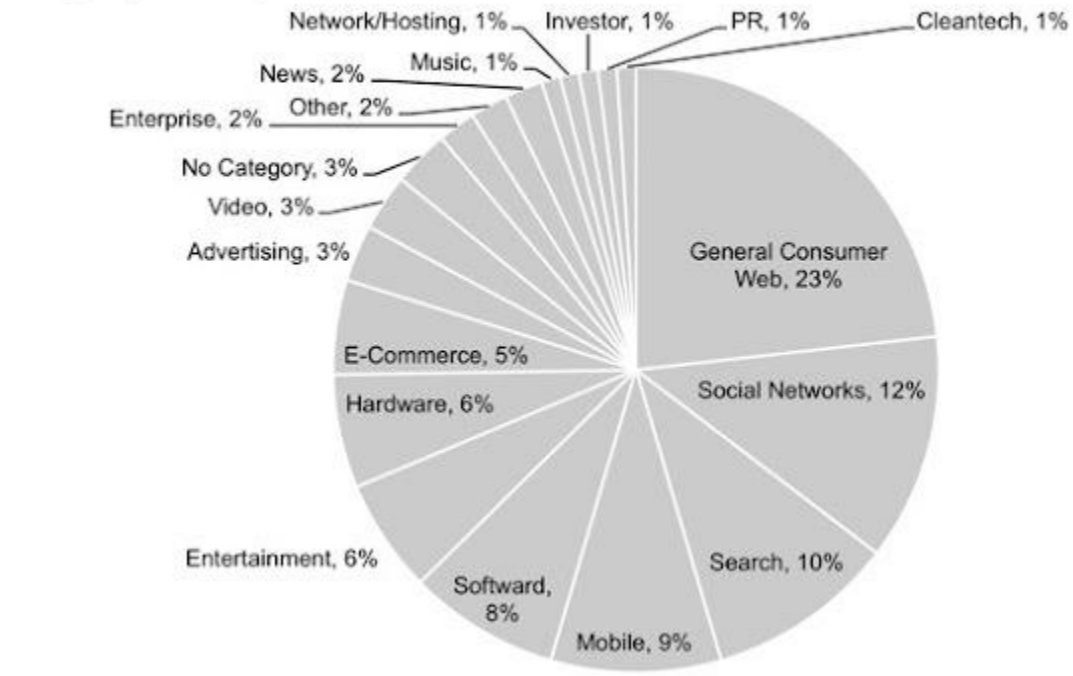


Share of coverage on TechCrunch

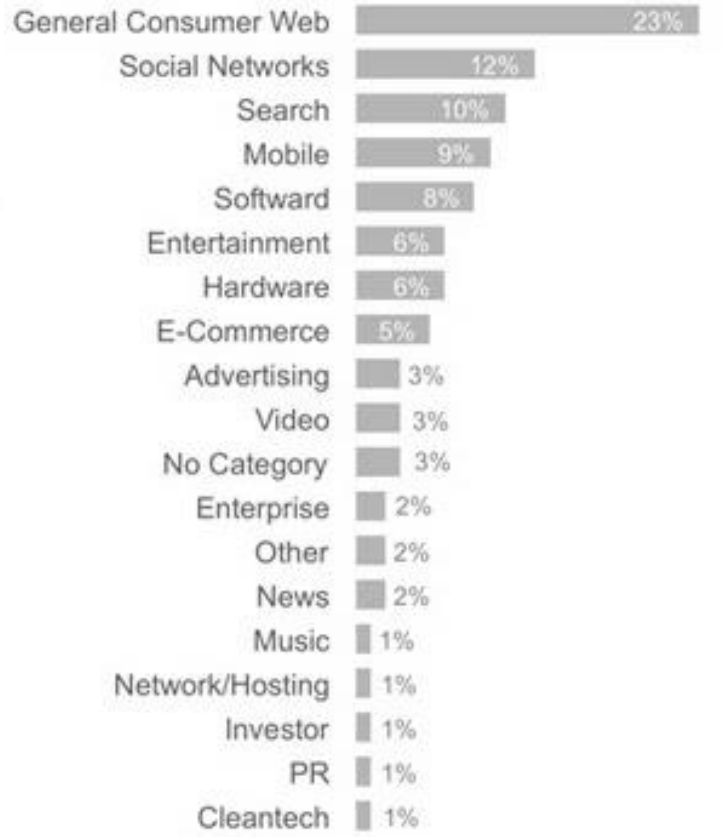
www.storytellingwithdata.com/2011/07/death-to-pie-charts.html
Cole Nussbaume

REDESIGN

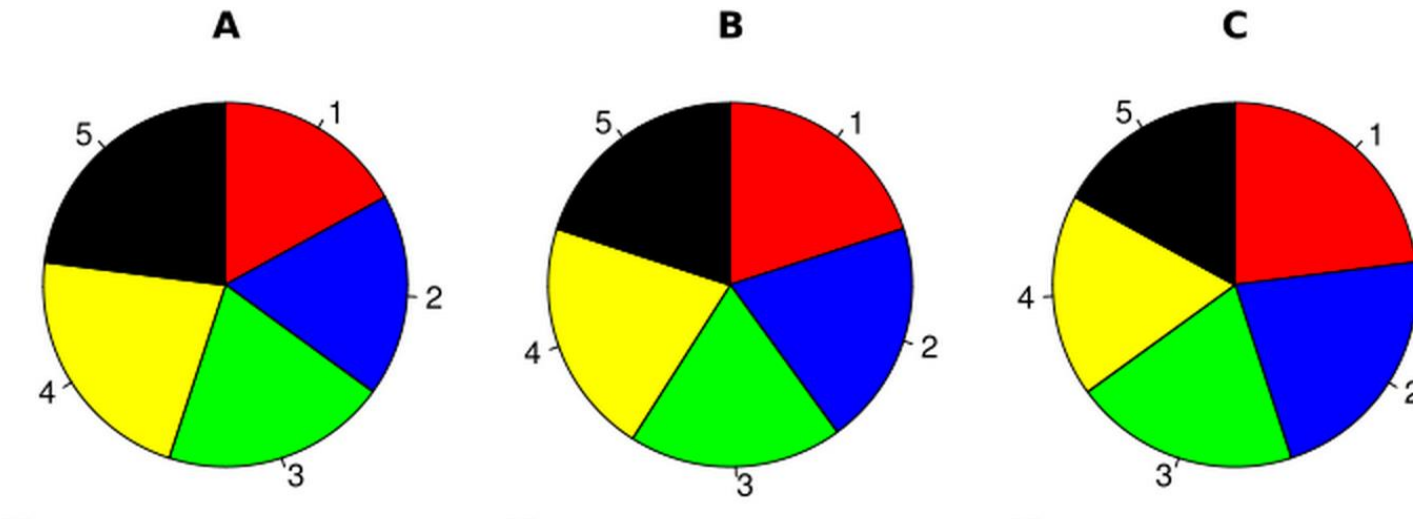
TechCrunch Coverage: 2005 - 2011
A slightly better pie?



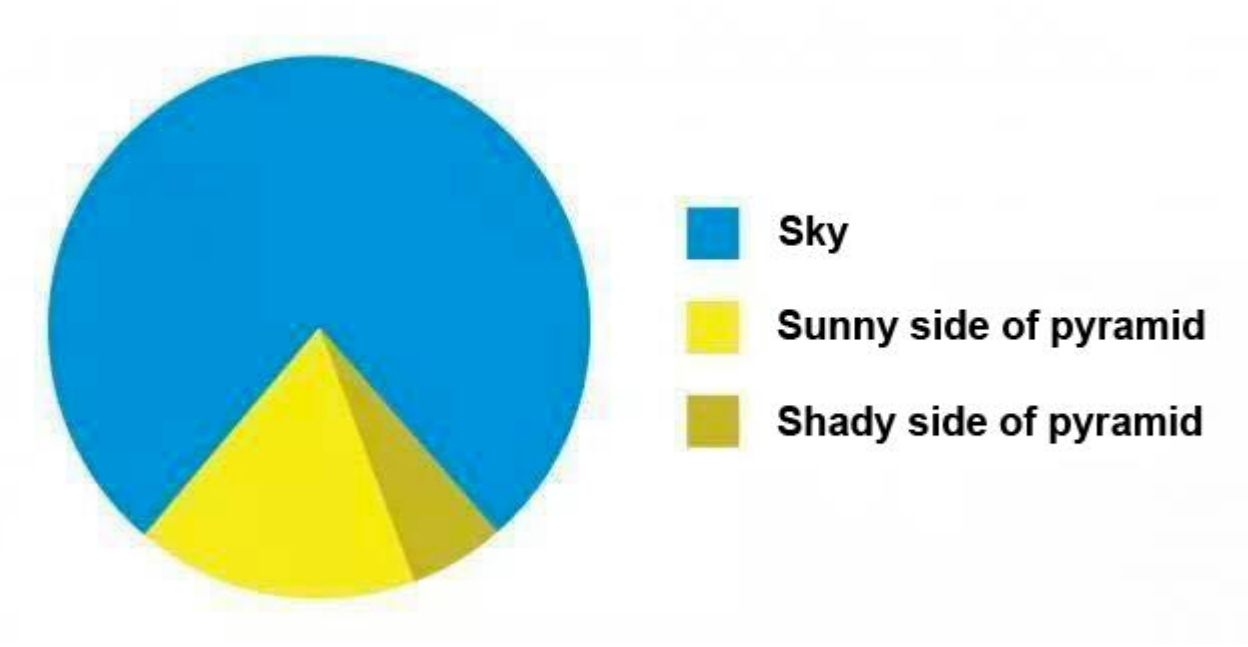
TechCrunch Coverage: 2005 - 2011
Bars are best!



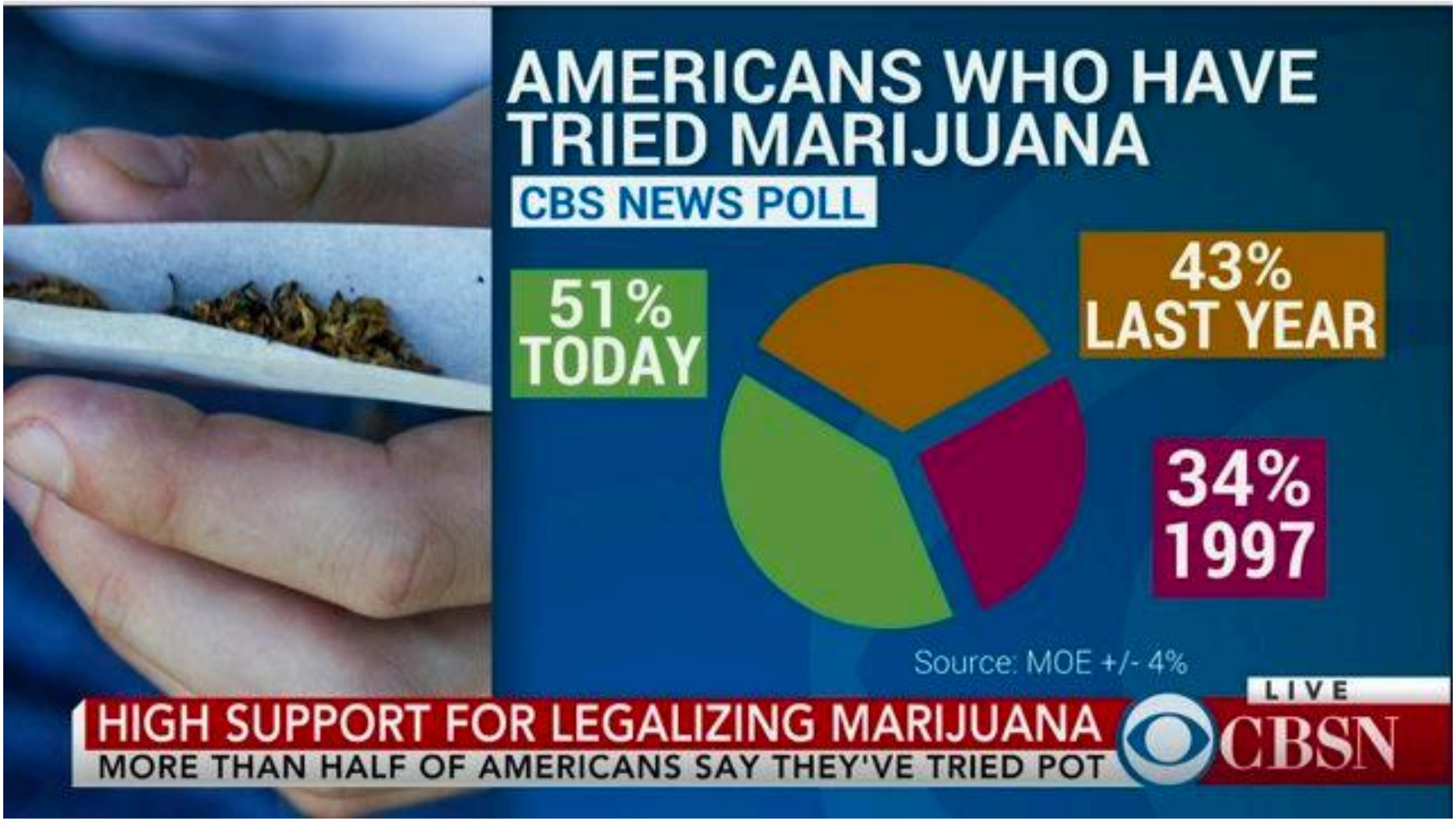
CAN YOU SPOT THE DIFFERENCES?



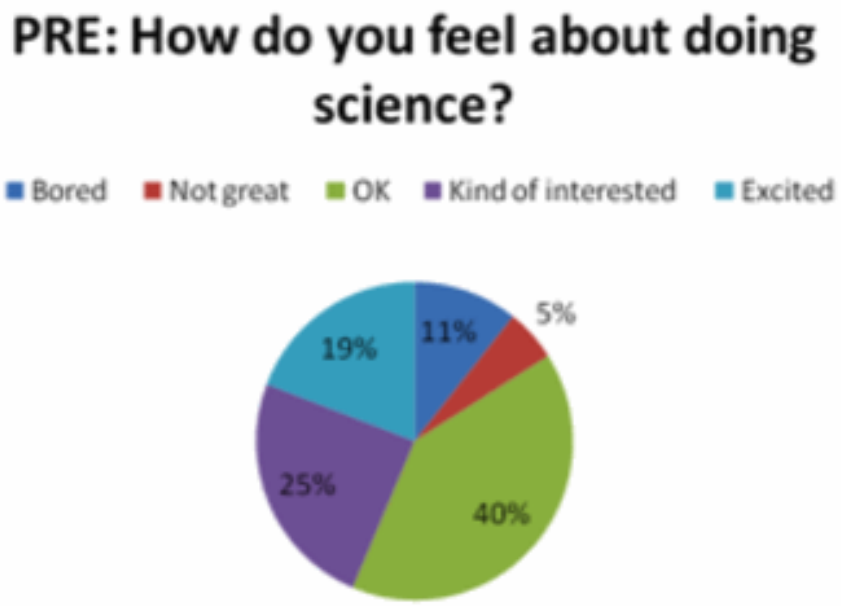
A GOOD PIE CHART?



BAD EXAMPLE AGAIN



SO, WHAT TO USE INSTEAD



http://www.storytellingwithdata.com/blog/2014/06/alternatives-to-pies

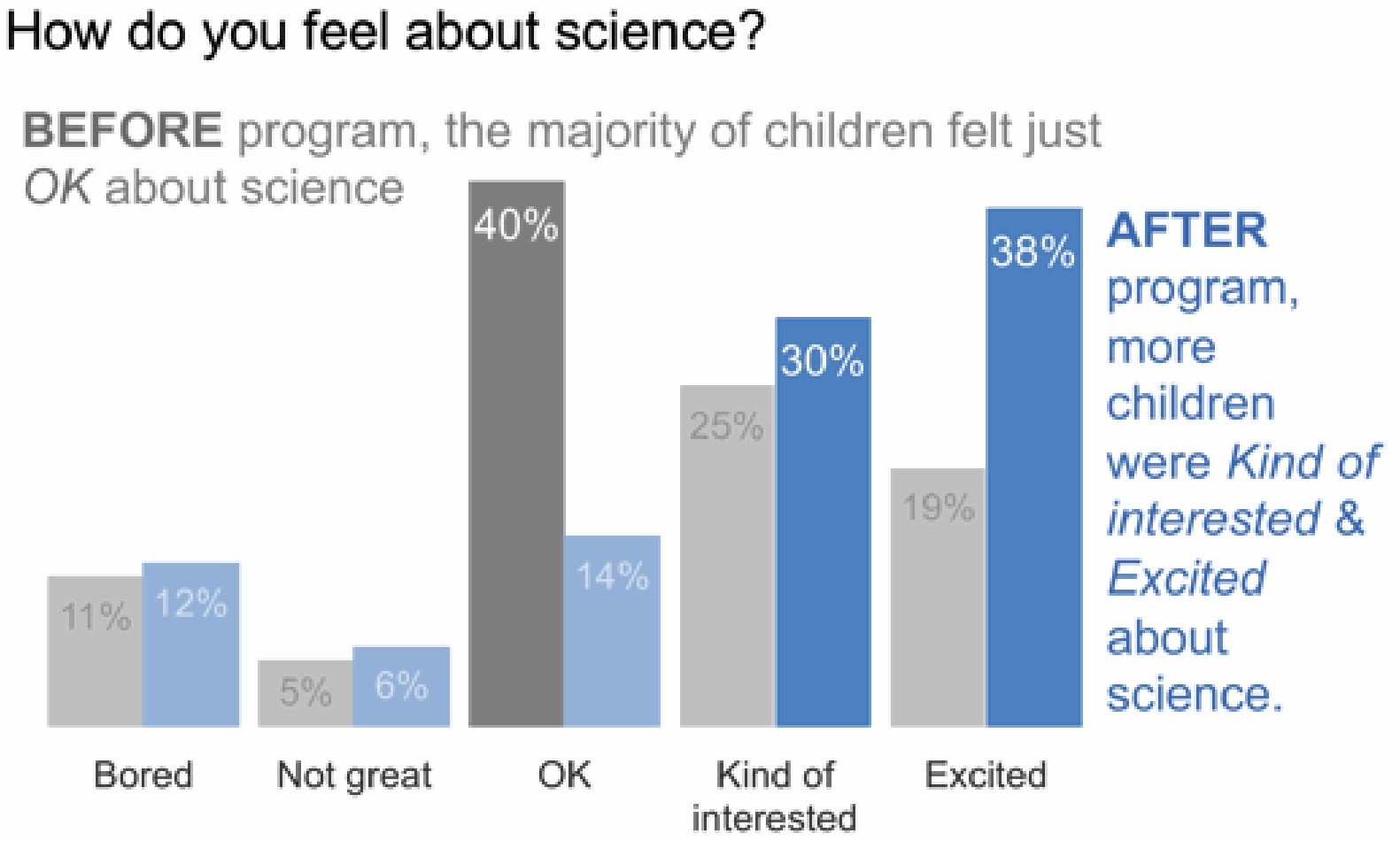
ALTERNATIVE #1. NUMBER DIRECTLY

After the pilot program,

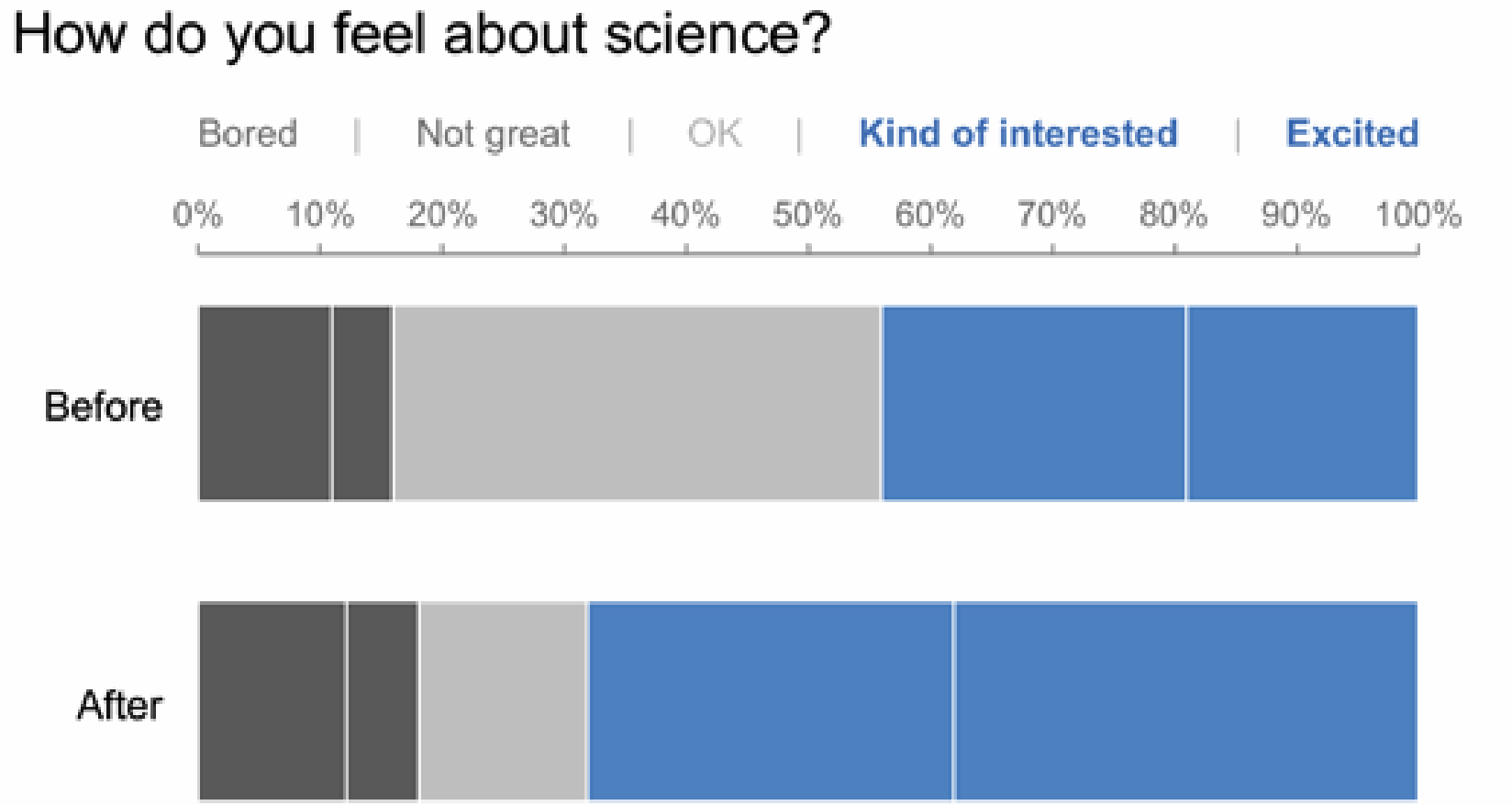
68%

of kids expressed interest towards science,
compared to 44% going into the program.

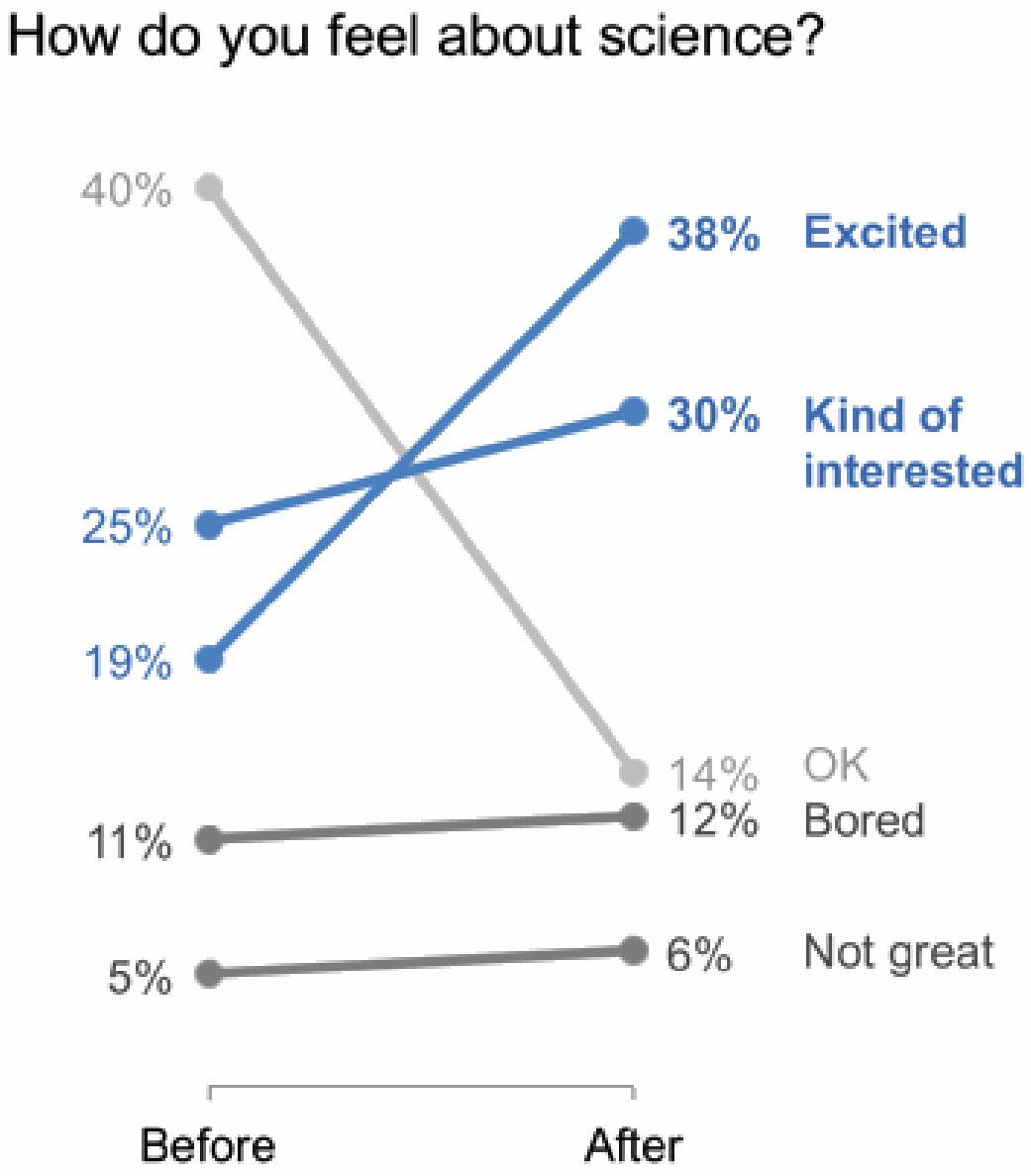
ALTERNATIVE #2: SIMPLE BAR GRAPH



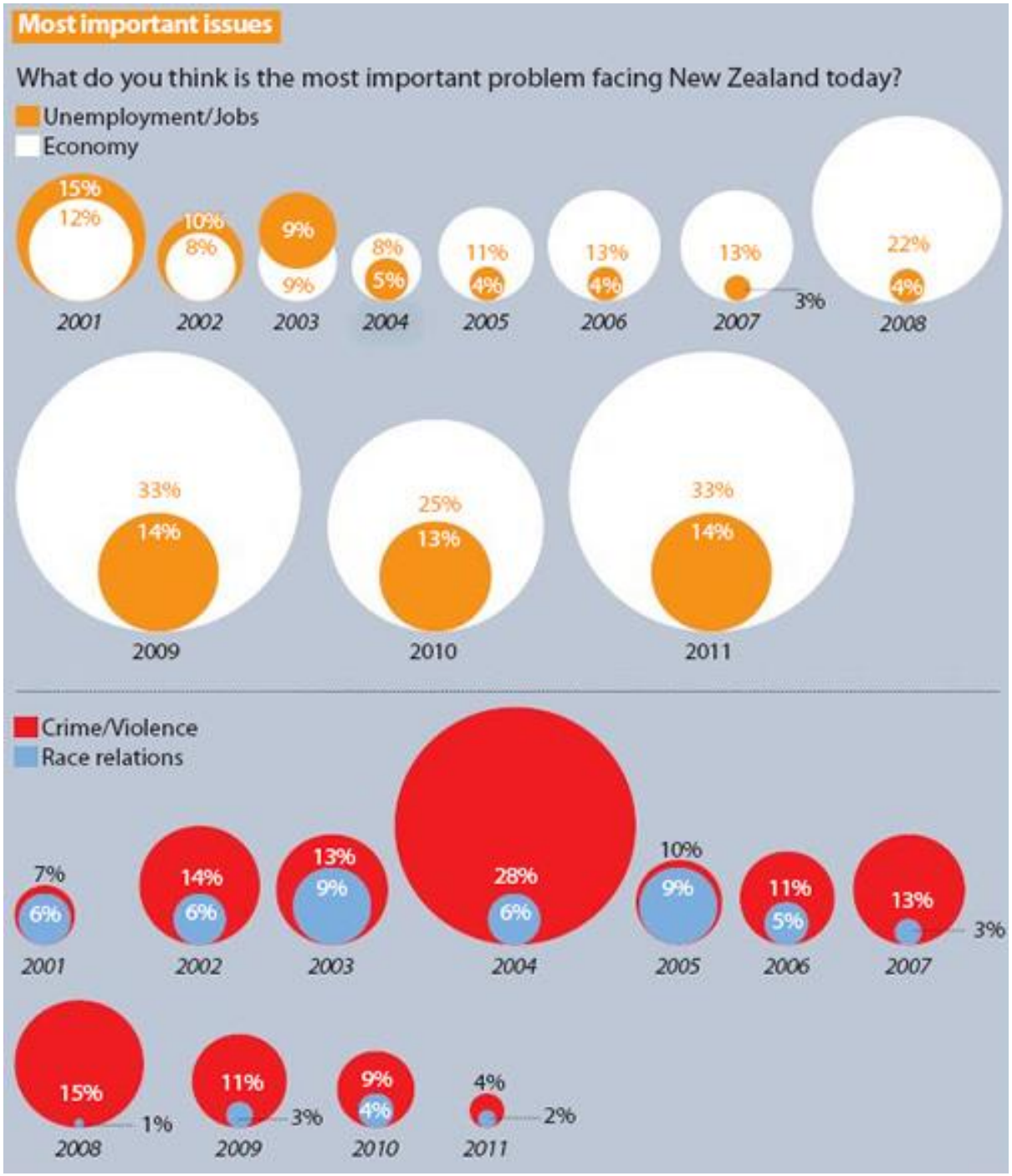
ALTERNATIVE #3: 100% STACKED HORIZONTAL BAR GRAPH

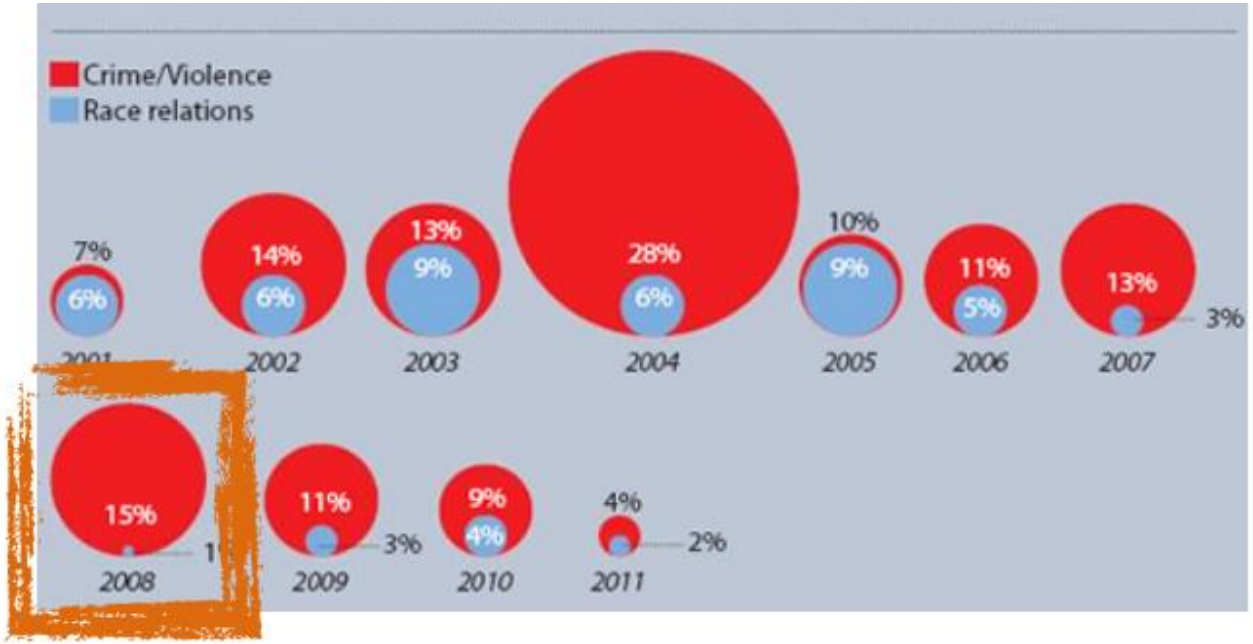


ALTERNATIVE #4: SLOPEGRAPH



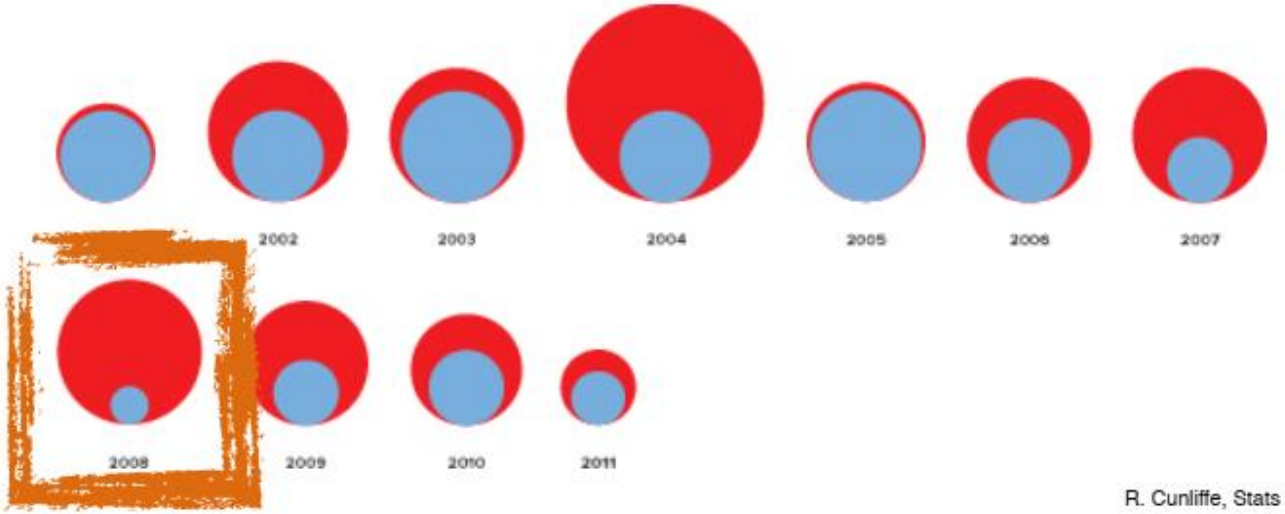
DESIGN CRITIQUE/ REDESIGN


Most important issues
What do you think is the most important problem facing New Zealand today?
Unemployment/Jobs
Economy
15%
12%
2001
10%
8%
2002
9%
9%
2003
8%
5%
2004
11%
4%
2005
13%
4%
2006
13%
3%
2007
22%
4%
2008
33%
14%
2009
25%
13%
2010
33%
14%
2011
Crime/Violence
Race relations
7%
6%
2001
14%
6%
2002
13%
9%
2003
28%
6%
2004
10%
9%
2005
11%
5%
2006
13%
3%
2007
15%
1%
2008
11%
3%
2009
9%
4%
2010
4%
2%
2011

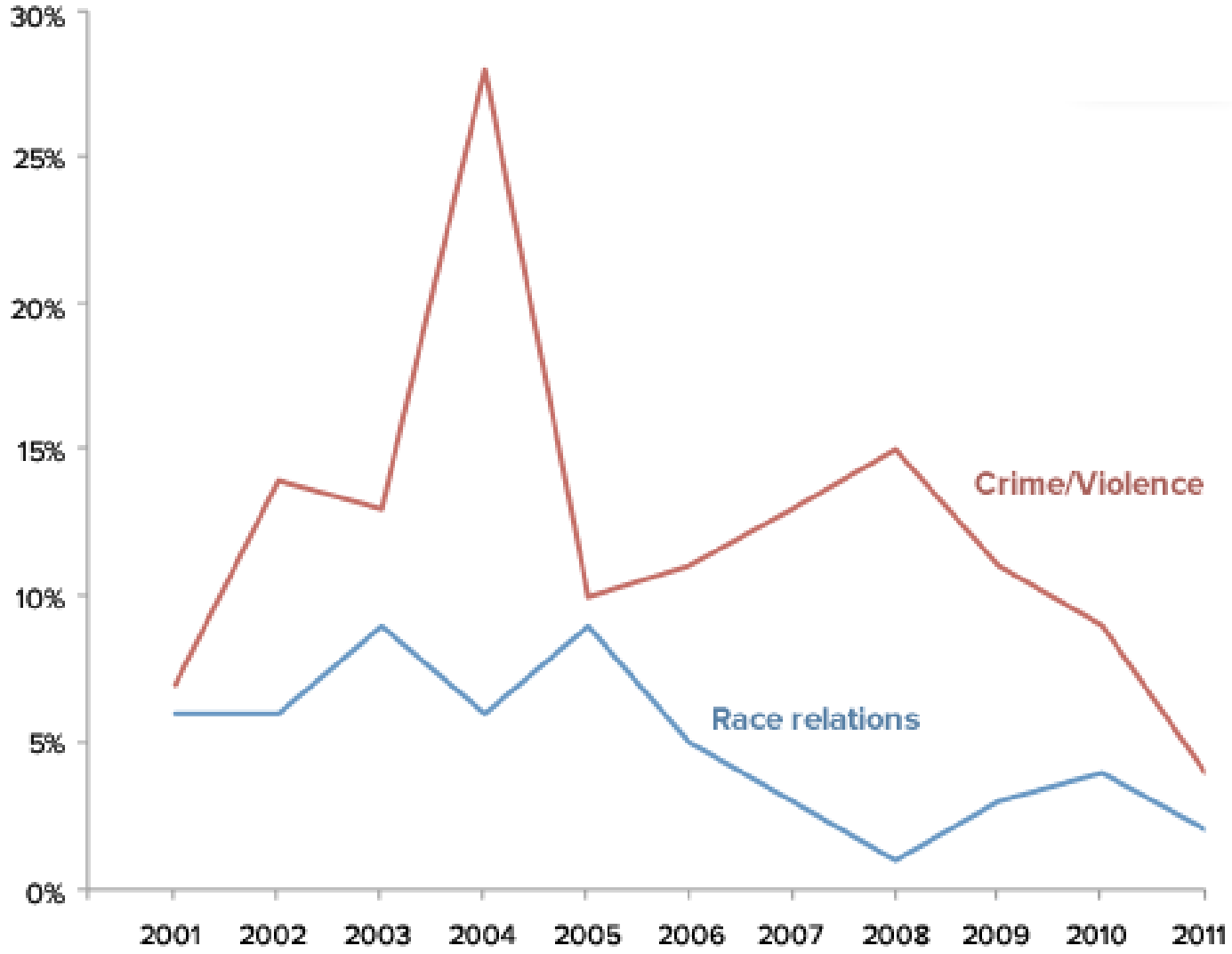


Quantity encoded by diameter, not area!

Fixing that:

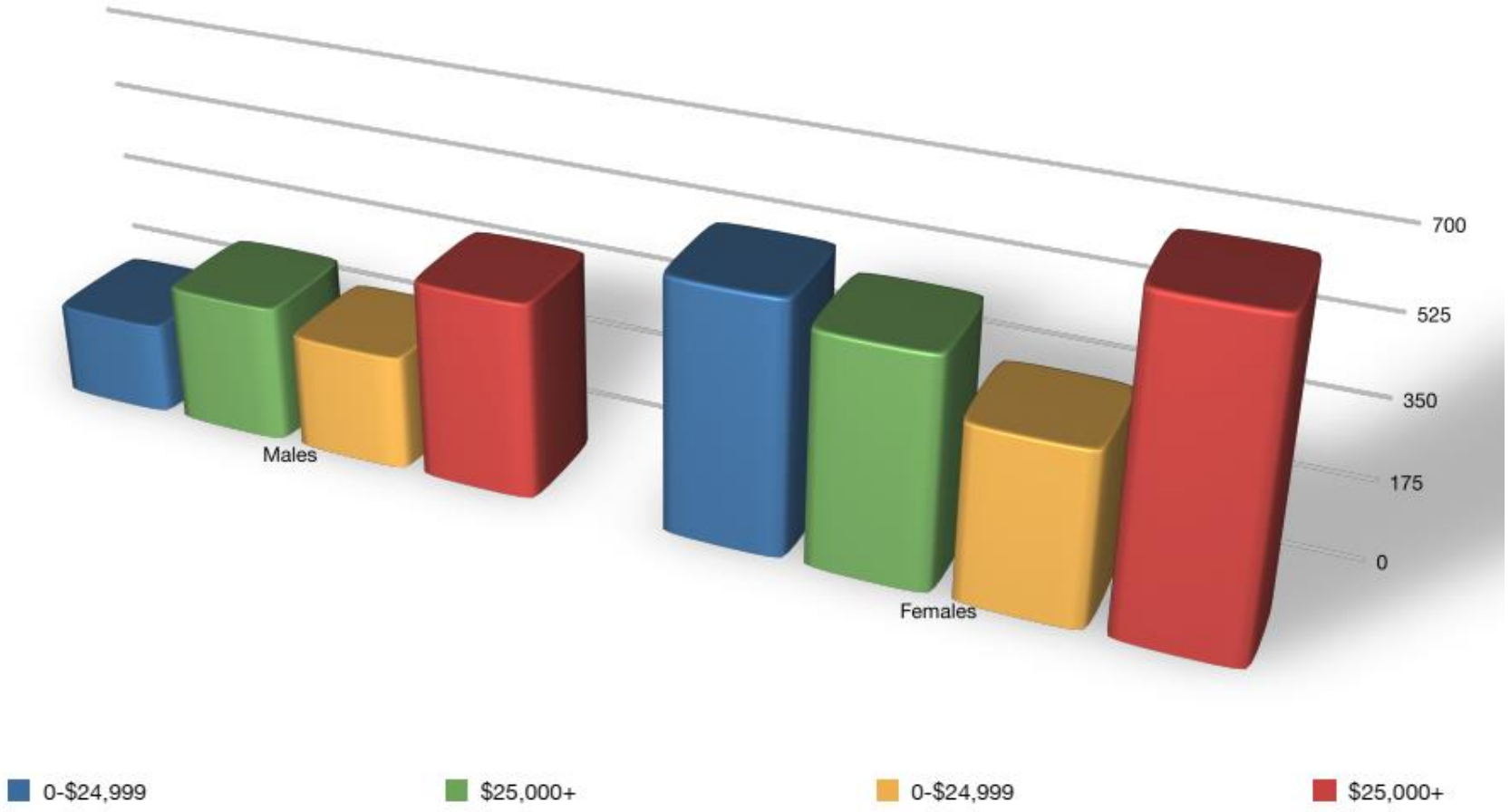


R. Cunliffe, Stats Chat

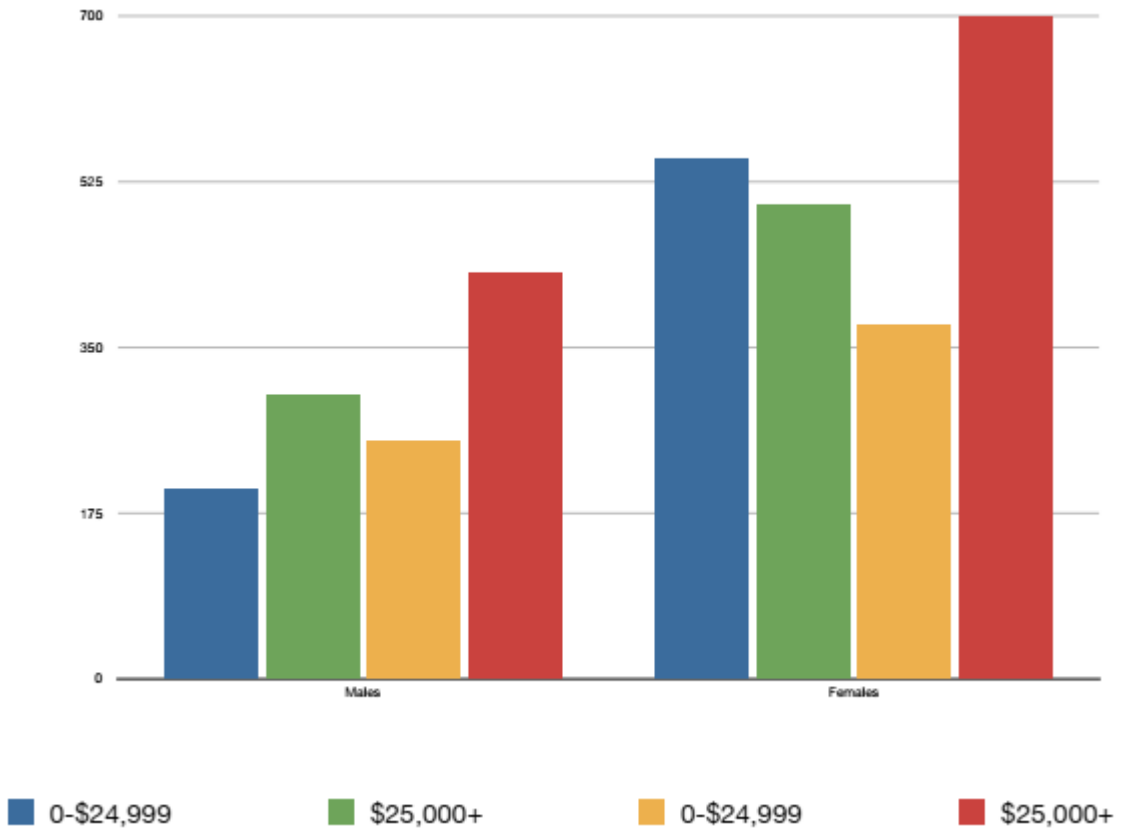

30%
25%
20%
15%
10%
5%
0%
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Crime/Violence
Race relations

VISUALIZATION DESIGN PRINCIPLES

MAXIMIZE DATA-INK RATIO

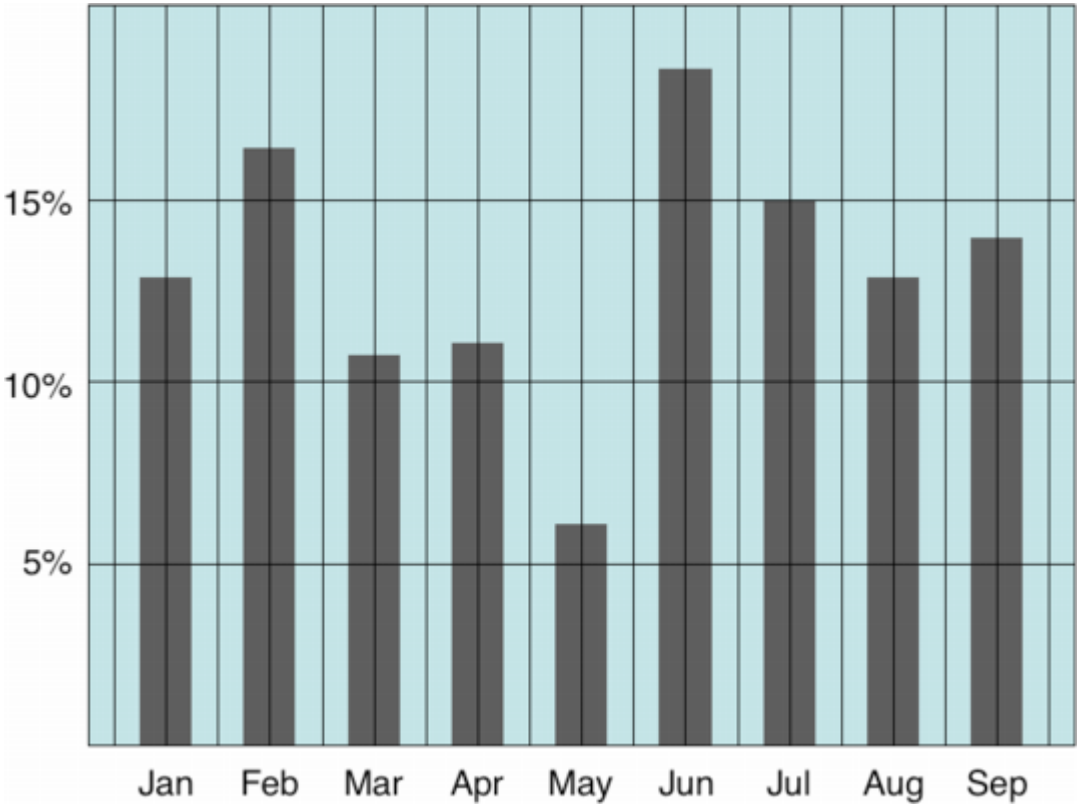


MAXIMIZE DATA-INK RATIO



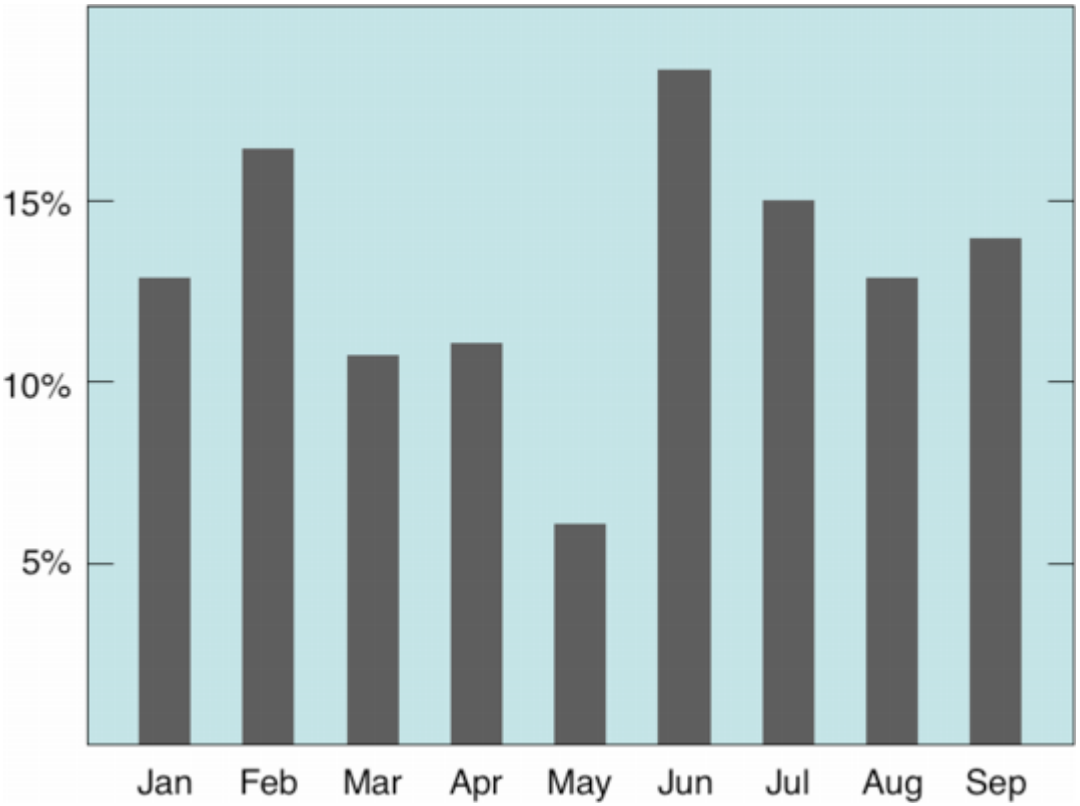
AVOID CHART JUNK

Extraneous visual elements that distract from the message

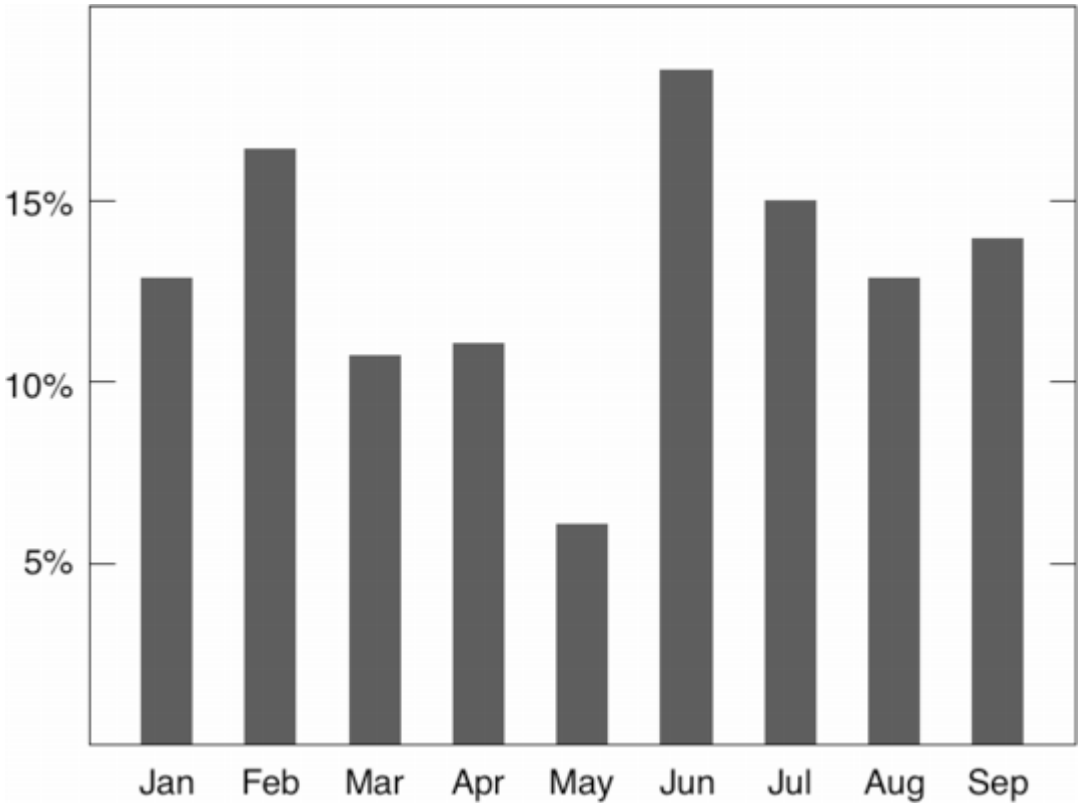


ongoing, Tim Brey

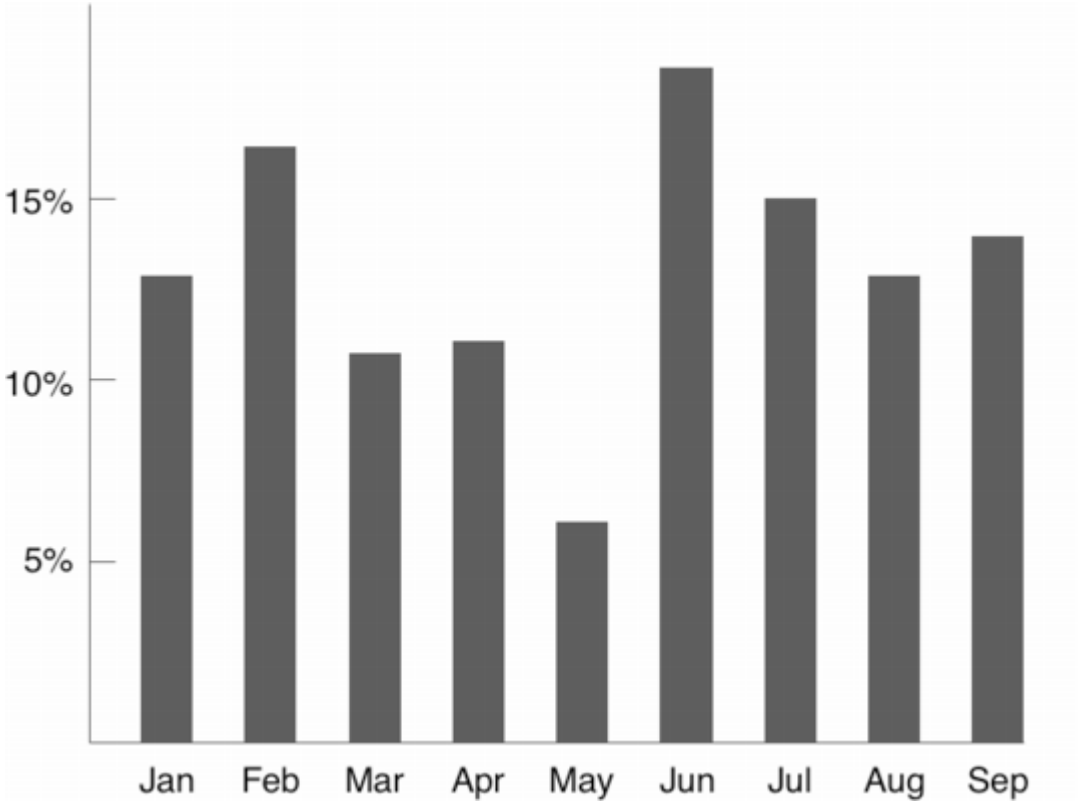
AVOID CHART JUNK



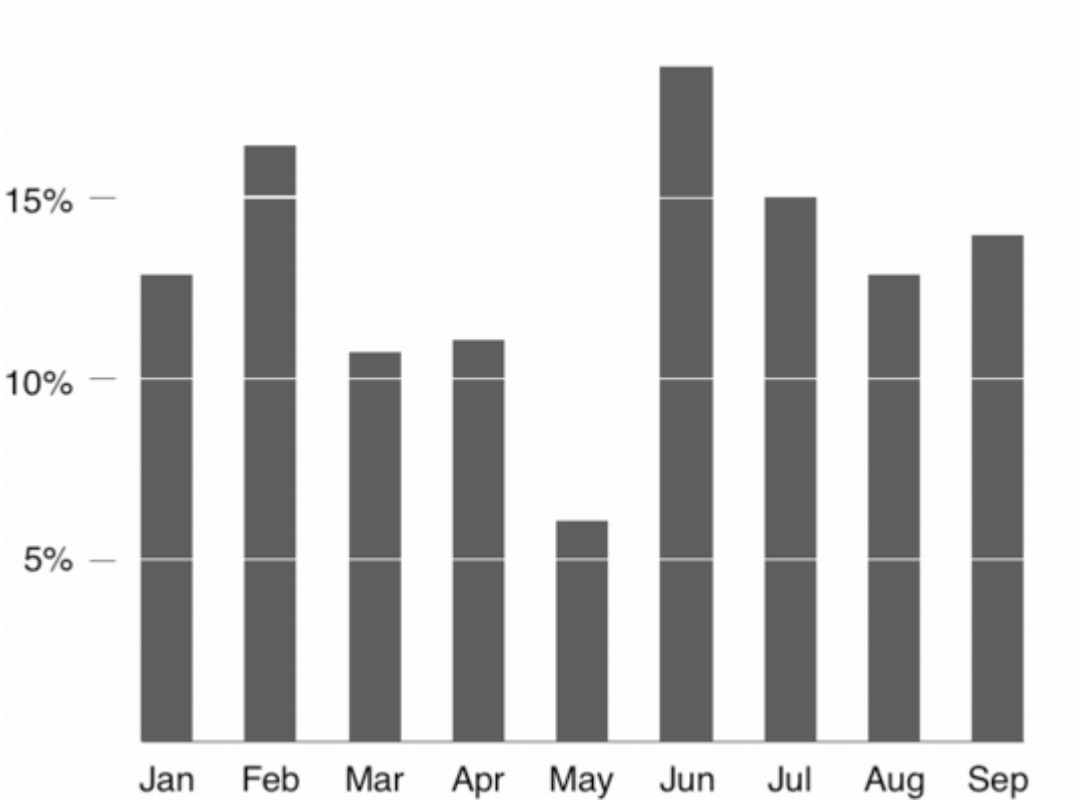
AVOID CHART JUNK



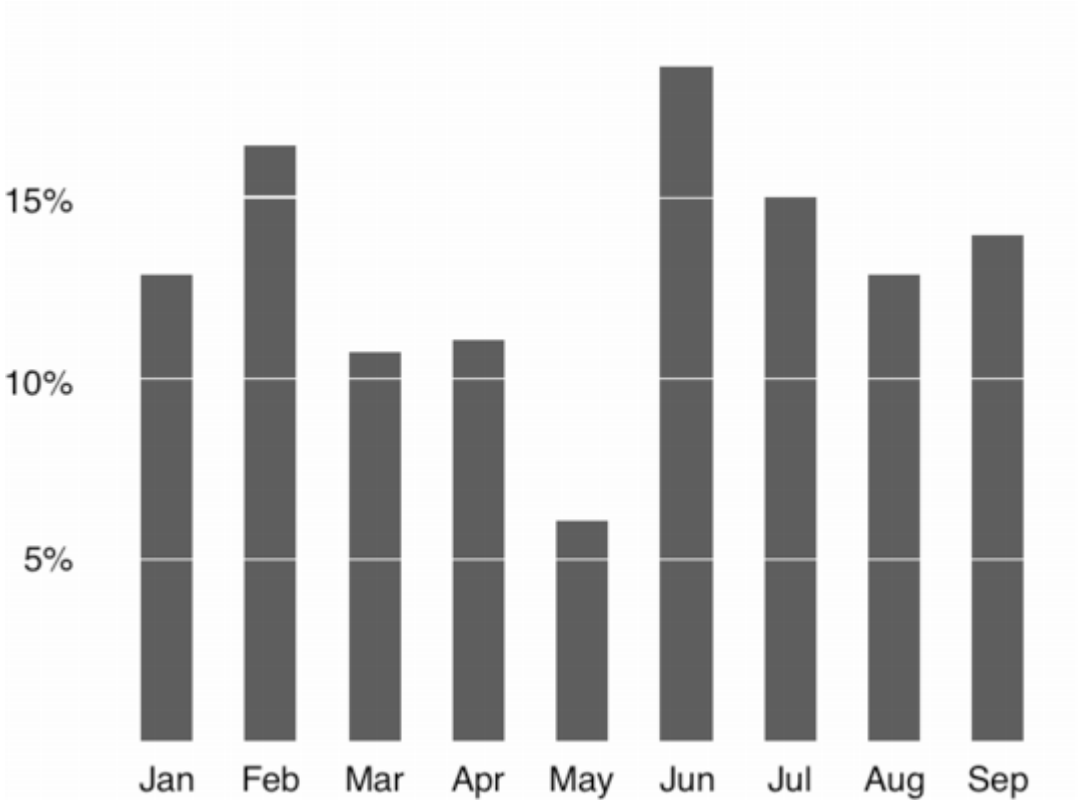
AVOID CHART JUNK



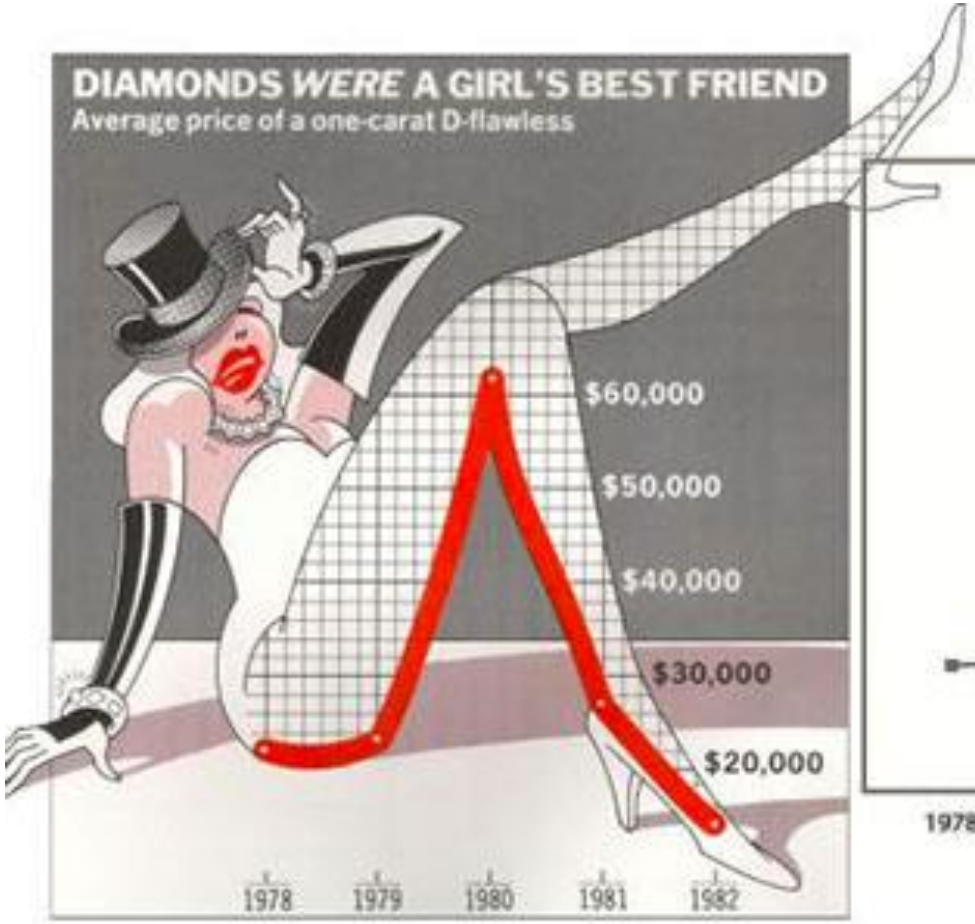
AVOID CHART JUNK

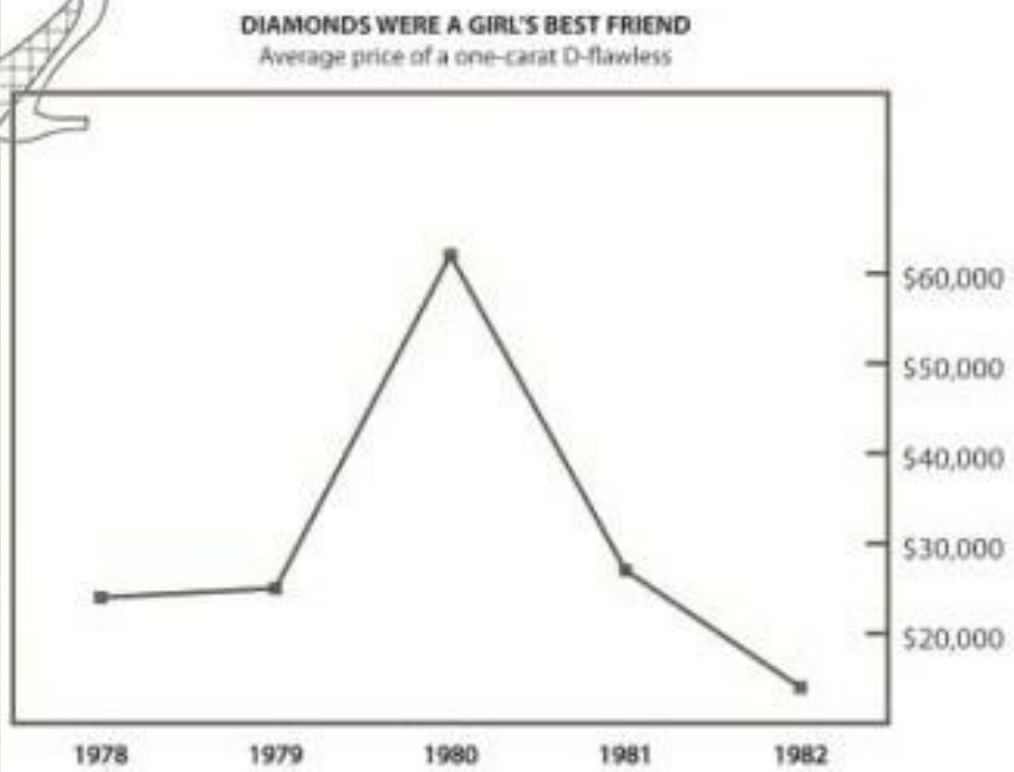


AVOID CHART JUNK



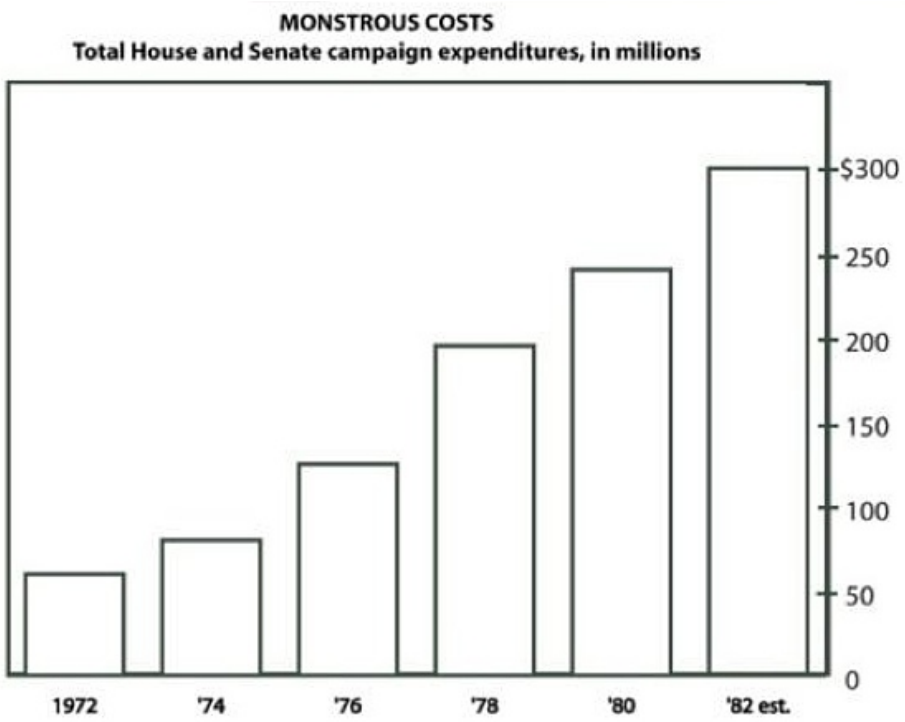
WHICH IS BETTER?





Bateman et al. 2010

WHICH IS BETTER?



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.

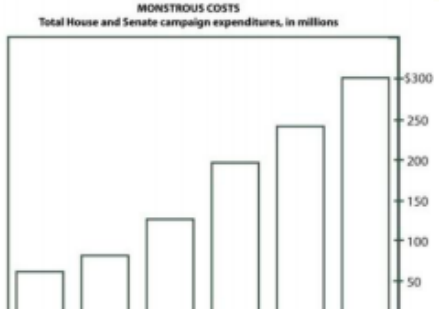
Author Keywords
Charts, information visualization, imagery, memorability.

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

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 *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).



EXPERIMENTAL RESULTS

1. No difference for **interpretation accuracy**
2. No difference in **recall accuracy after a five-minute gap**
3. Significantly **better recall for Holmes charts** of both the chart topic and the details (categories and trend) **after long-term gap** (2-3 weeks).
4. Participants **saw value messages** in the Holmes charts significantly more often than in the plain charts.
5. Participants found the Holmes charts **more attractive, most enjoyed** them, and found that they were **easiest and fastest to remember**.

USE CHART JUNK? IT DEPENDS!

PROS

Persuasion

Memorability

Engagement

CONS

Biased analysis

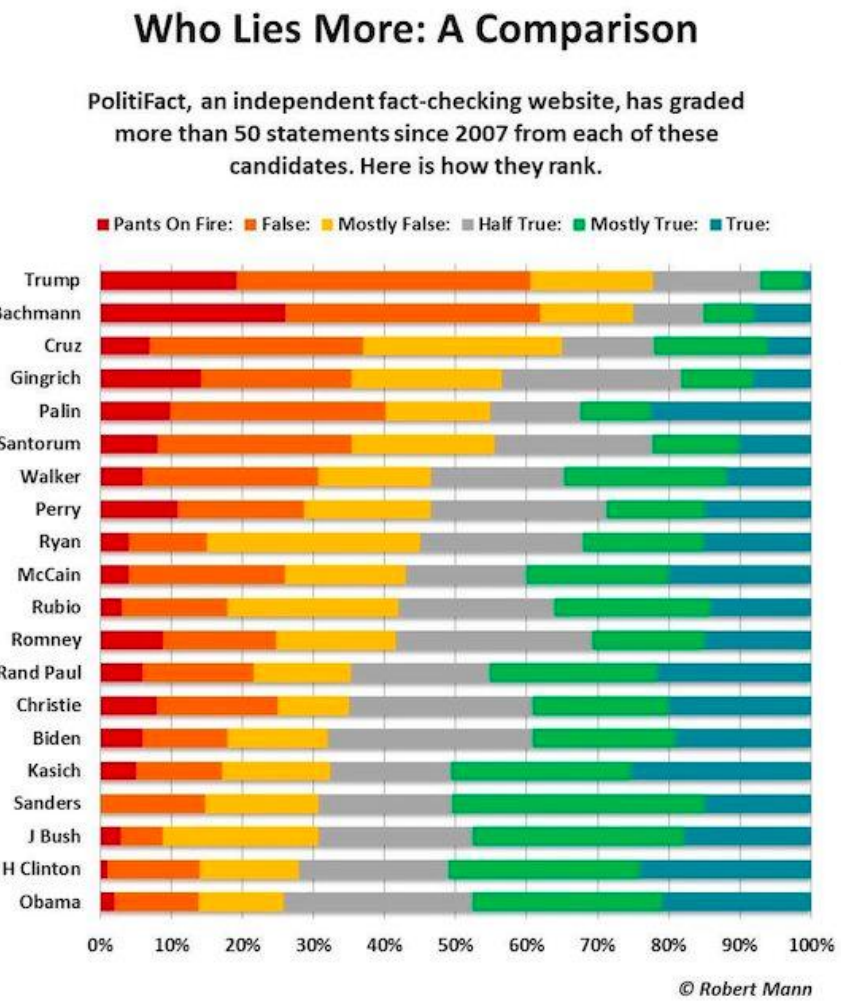
Trustworthiness

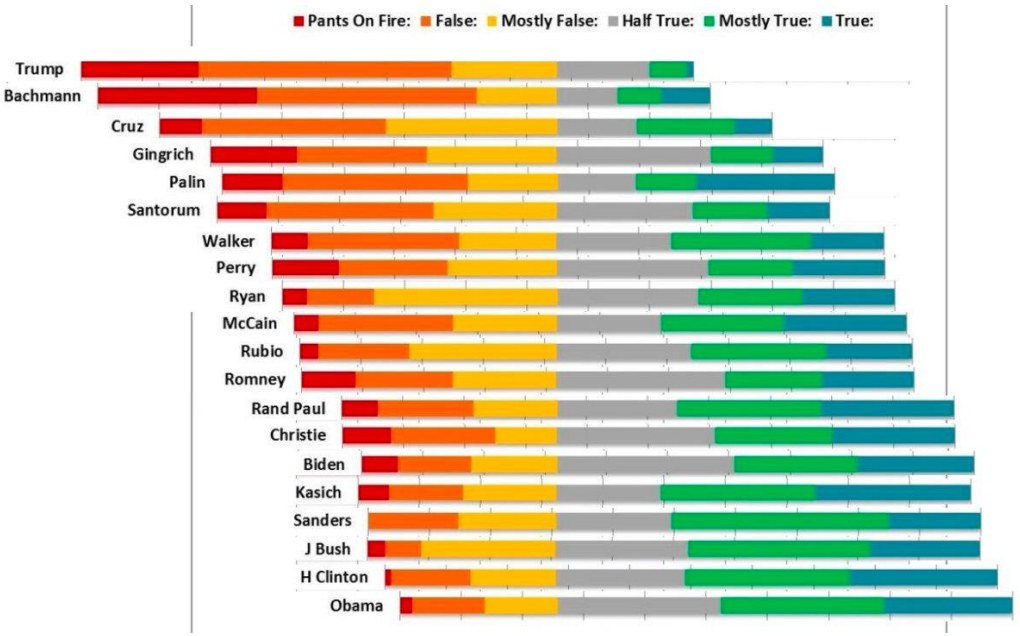
Interpretability

Space efficiency

Effort

ALIGNMENT MATTERS





https://twitter.com/infowetrust/status/760521739092627457

http://www.visualisingdata.com/2016/08/little-visualisation-design-part-21/

NO UNJUSTIFIED 3D

Depth judgment is bad

N = 0.67

Occlusion

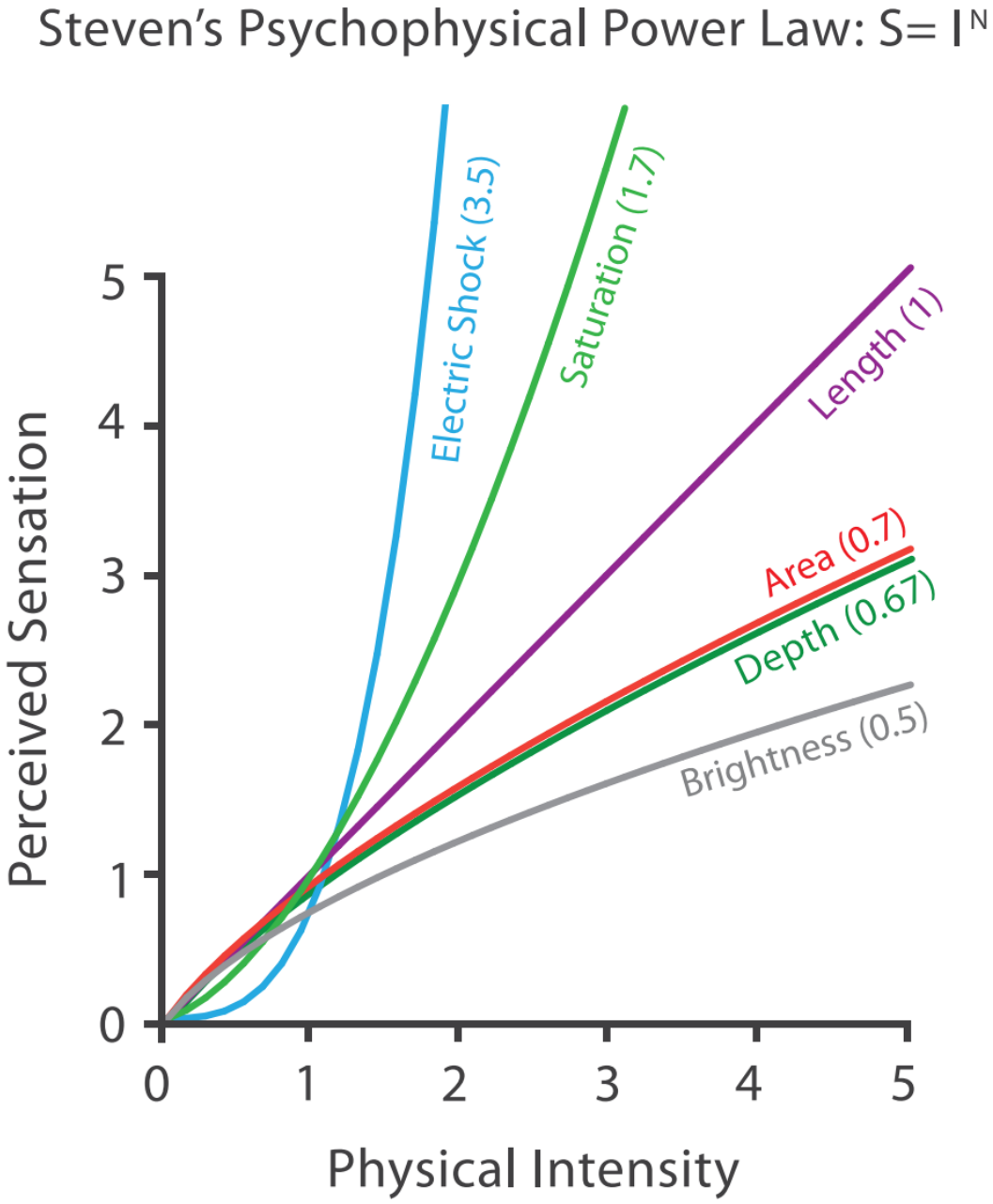
Perspective distortion

Color: Lighting /

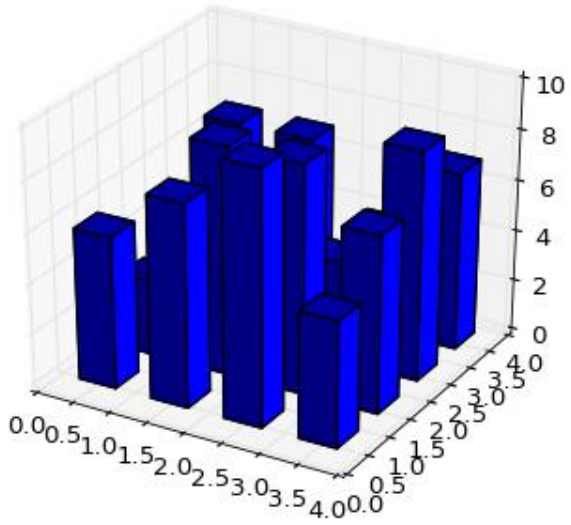
Shadows /

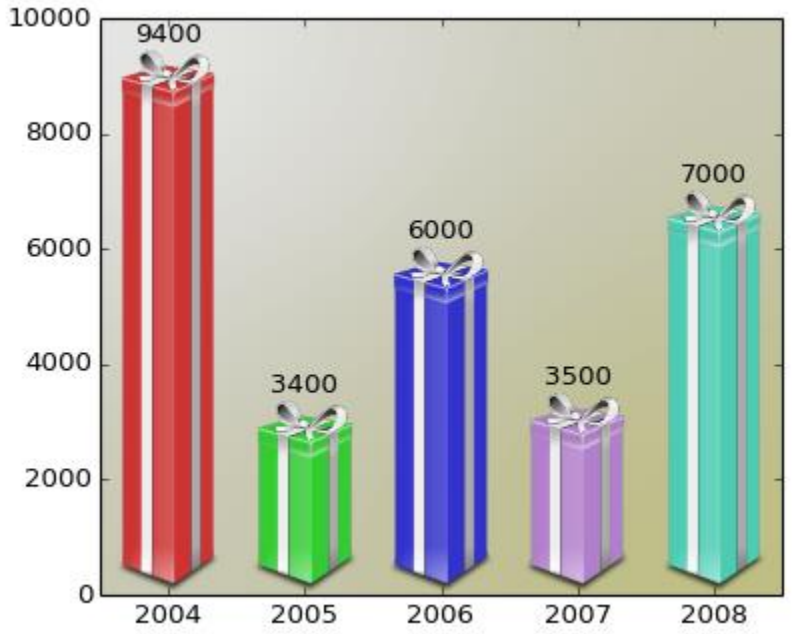
Shading

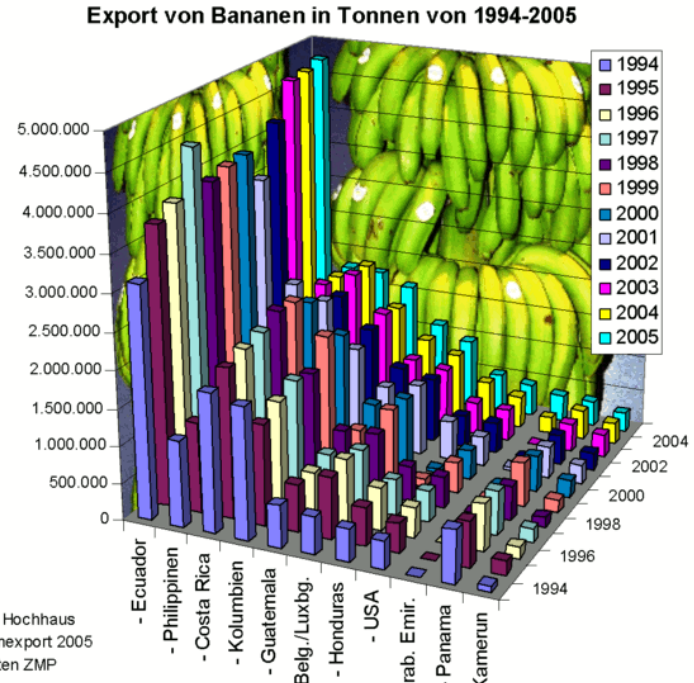
Title text illegible

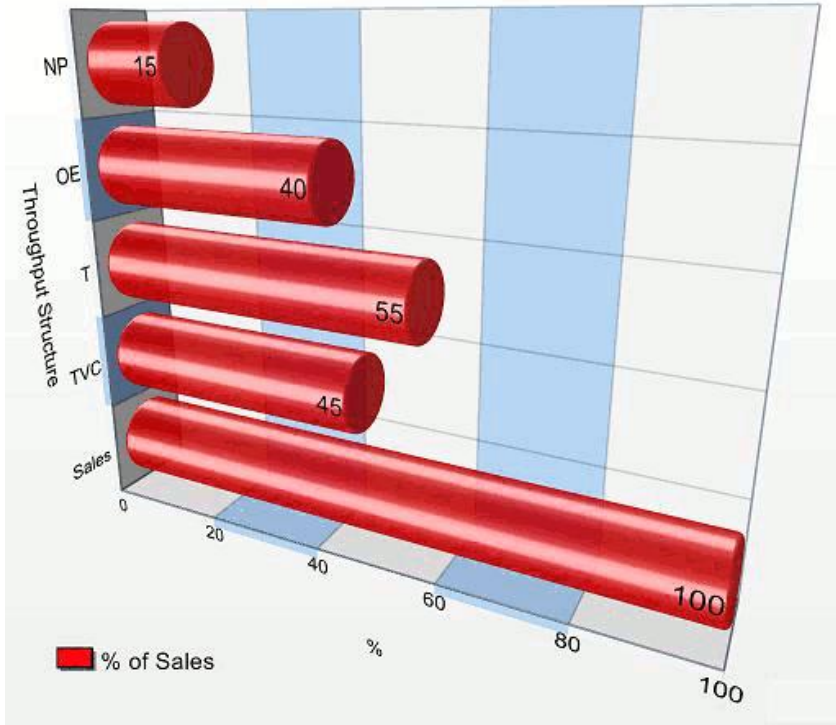


DON'T

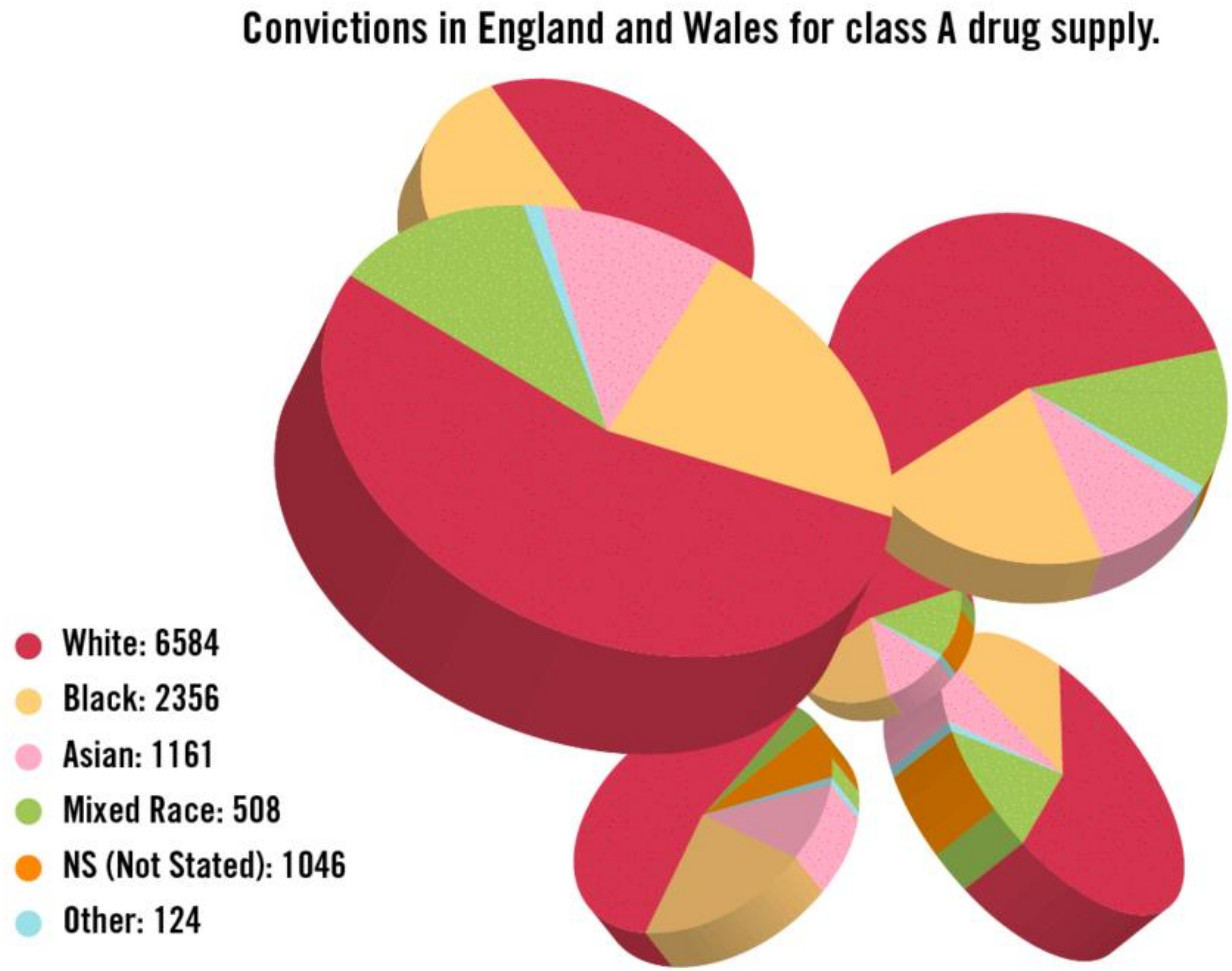

0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0
2
4
6
8
10


10000
8000
6000
4000
2000
0
9400
3400
6000
3500
7000
2004
2005
2006
2007
2008

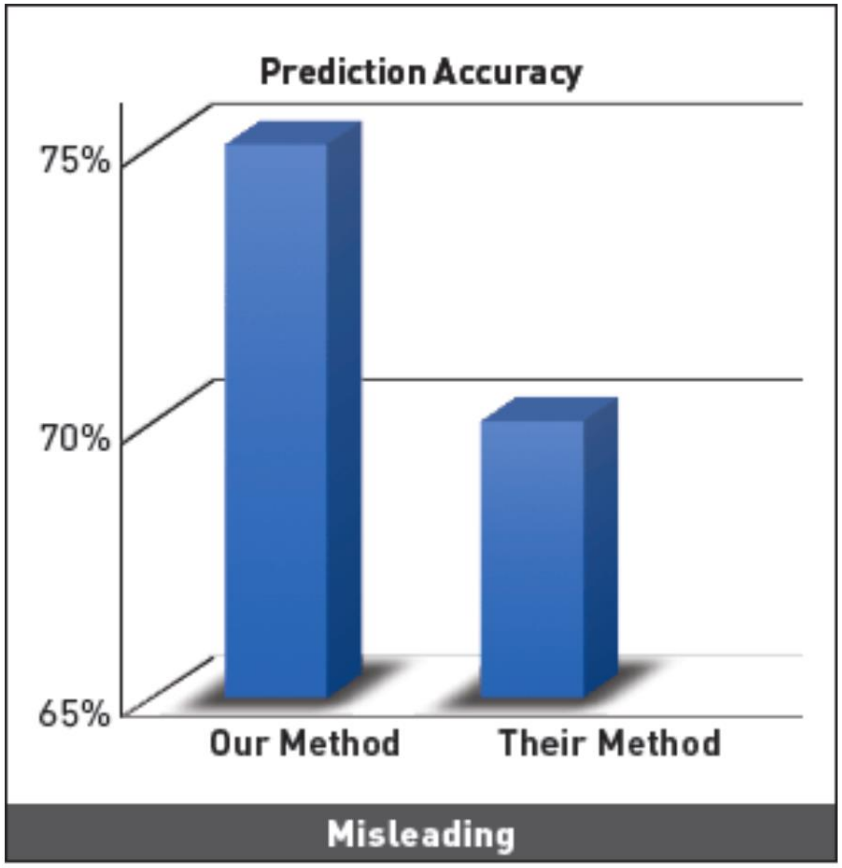

Export von Bananen in Tonnen von 1994-2005
5.000.000
4.500.000
4.000.000
3.500.000
3.000.000
2.500.000
2.000.000
1.500.000
1.000.000
500.000
0
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2004
2002
2000
1998
1996
1994
- Ecuador
- Philippinen
- Costa Rica
- Kolumbien
- Guatemala
- Belg./Luxbg.
- Honduras
- USA
rab. Emir.
Panama
Kamerun
Dr. Hochhaus
Banexport 2005
Daten ZMP

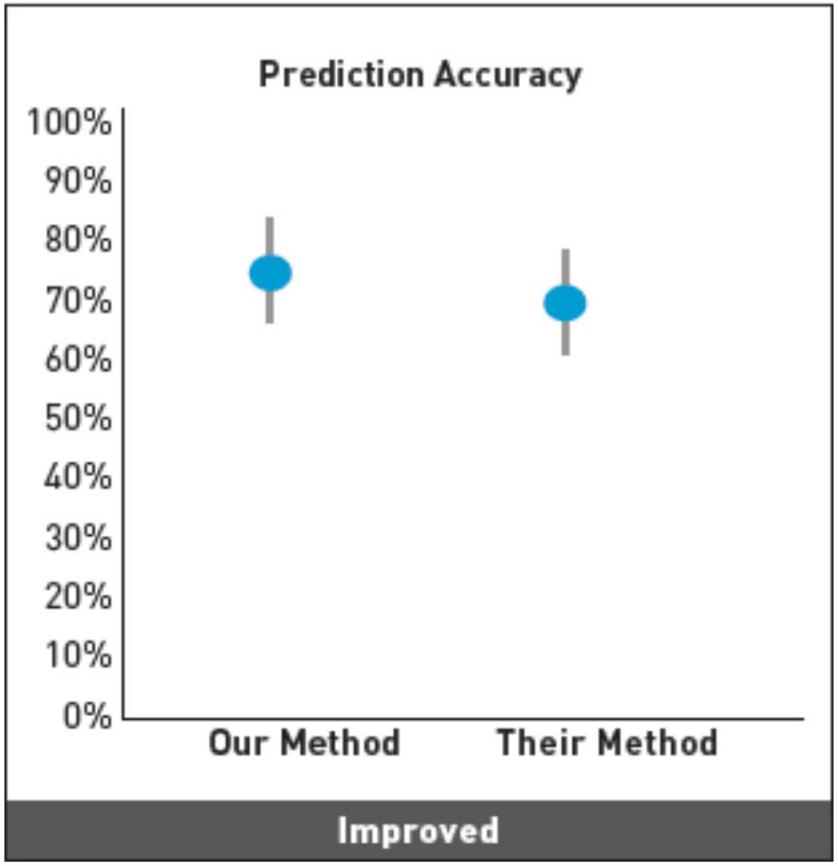

NP
15
OE
40
T
55
TVC
45
Sales
100
Throughput Structure
0
20
40
60
80
100
%
% of Sales

DON'T


Convictions in England and Wales for class A drug supply.
White: 6584
Black: 2356
Asian: 1161
Mixed Race: 508
NS (Not Stated): 1046
Other: 124

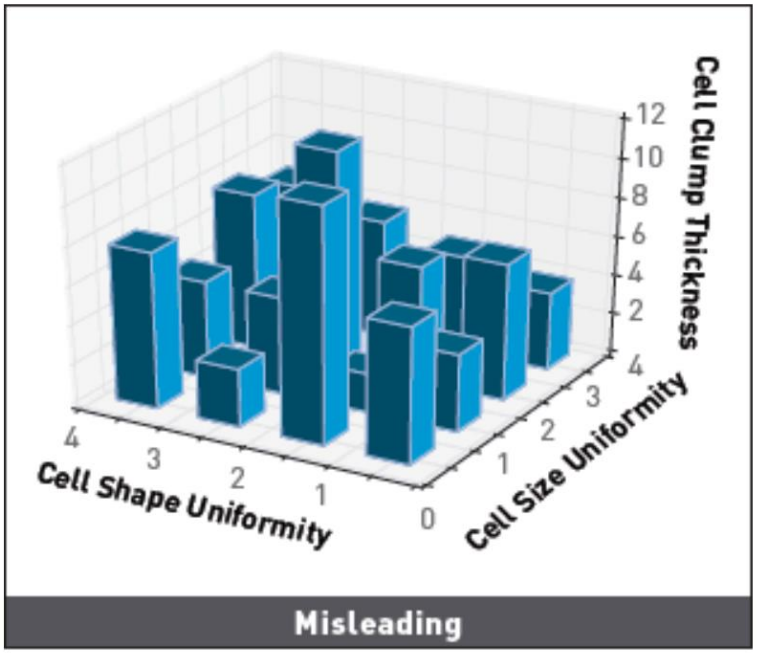
3D DESIGN ALTERNATIVE



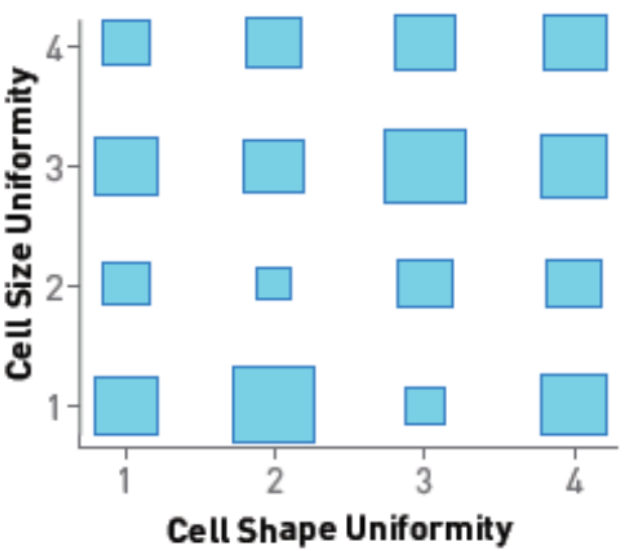


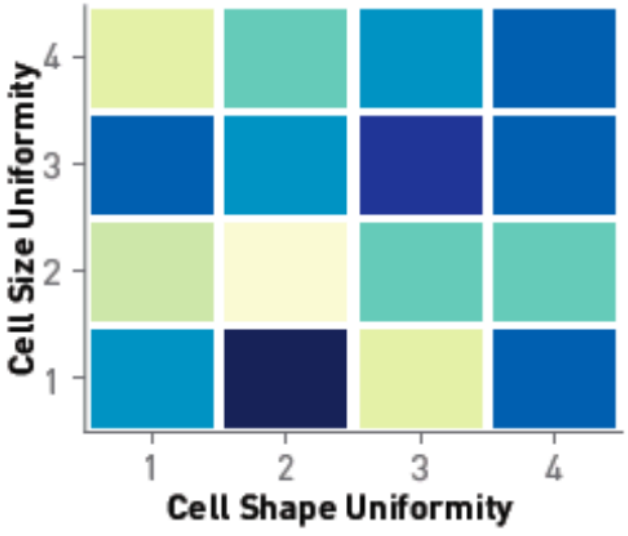
http://interactions.acm.org/archive/view/july-august-2018/the-good-the-bad-and-the-biased

3D DESIGN ALTERNATIVE



Misleading



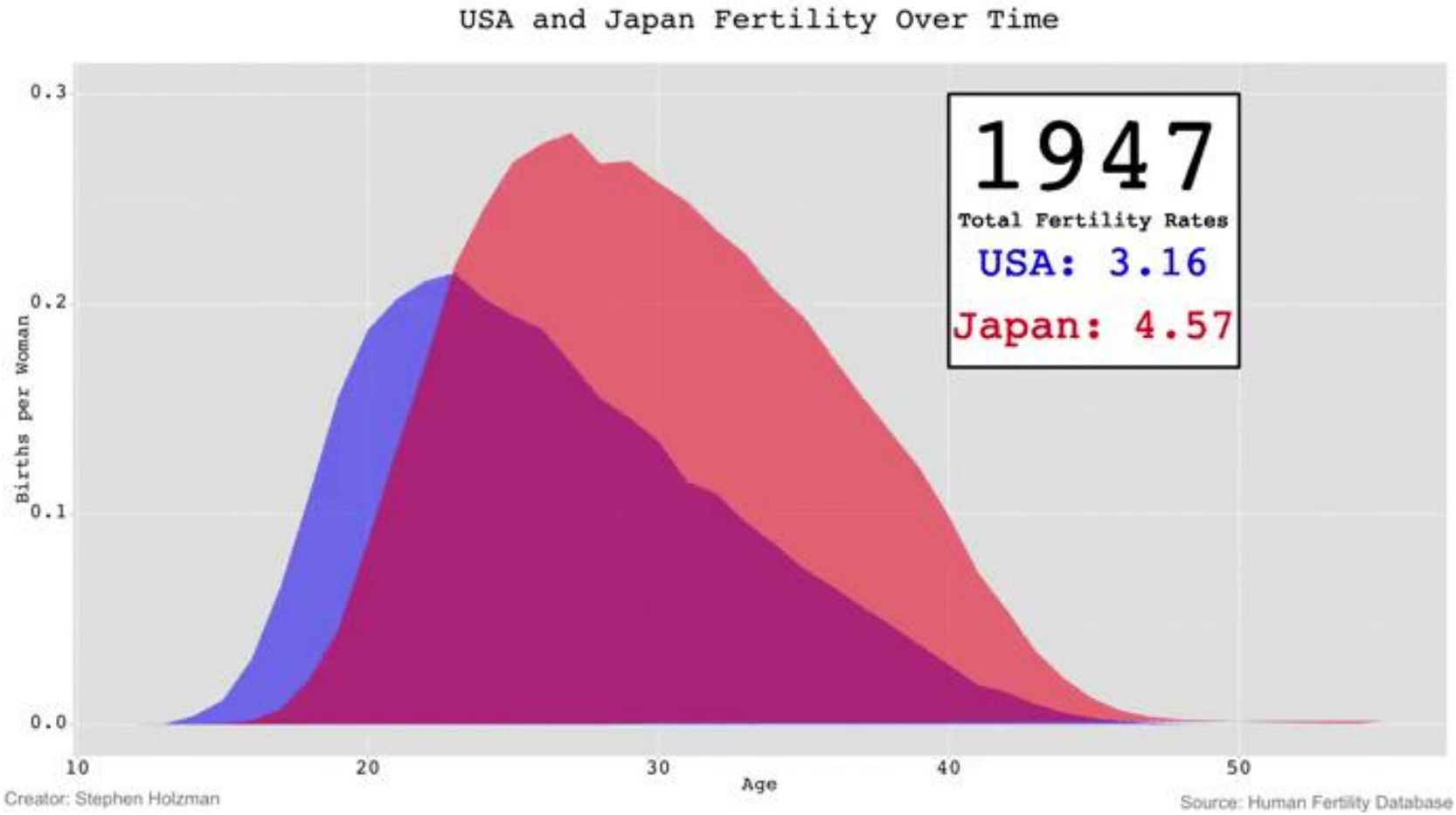


Improved

http://interactions.acm.org/archive/view/july-august-2018/the-good-the-bad-and-the-biased

EYES BEAT MEMORY

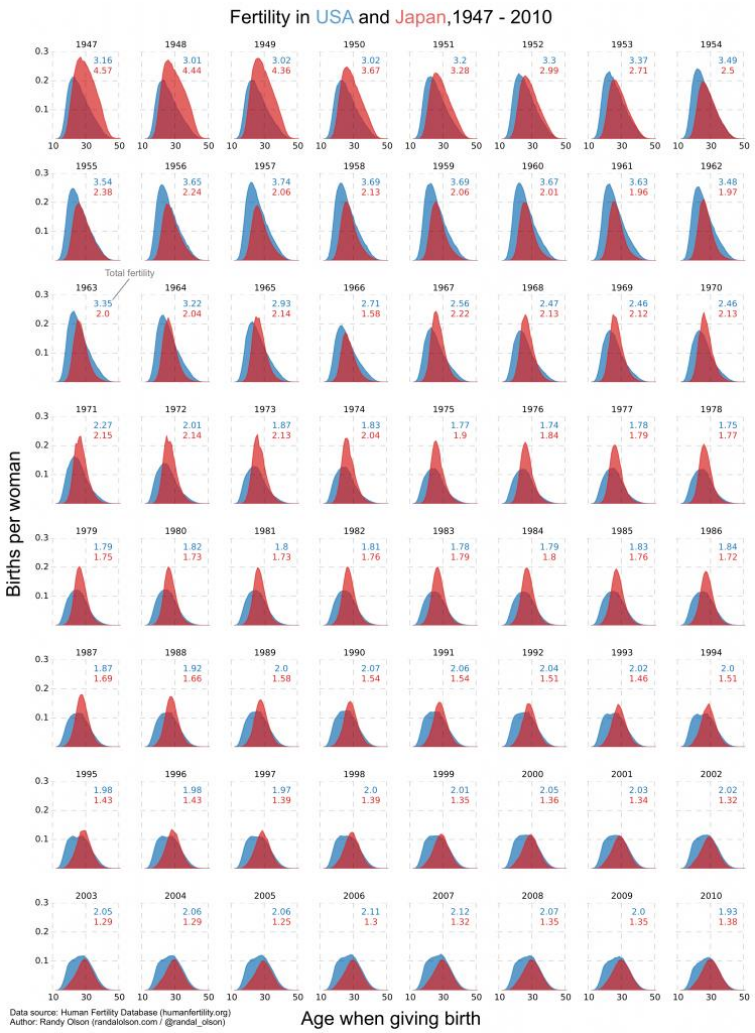
Don’t make people memory, show them



http://www.randalolson.com/2015/08/23/small-multiples-vs-animated-gifs-for-showing-changes-in-fertility-rates-over-time/

What can we do differently?

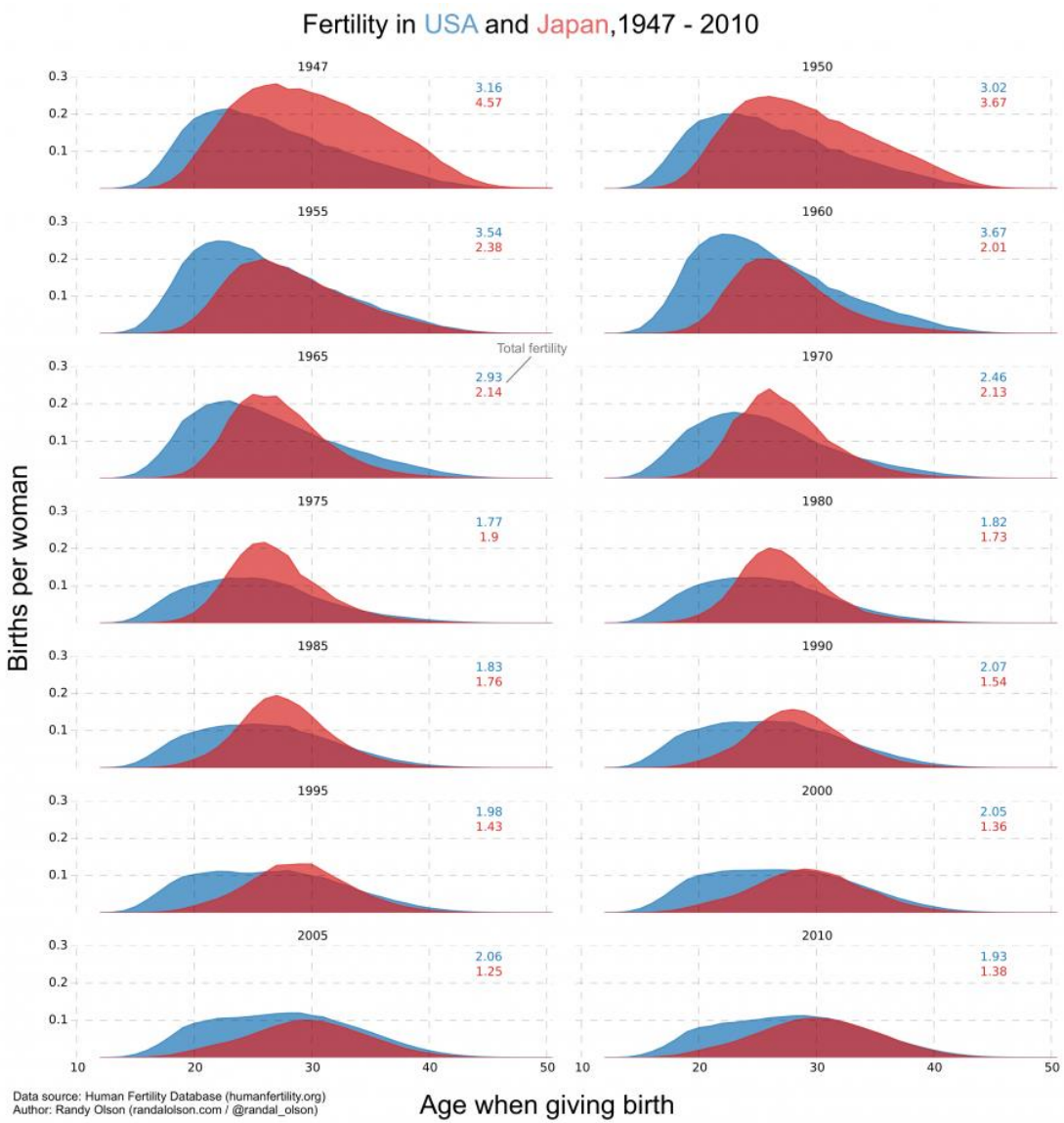
EYES BEAT MEMORY: SMALL MULTIPLES



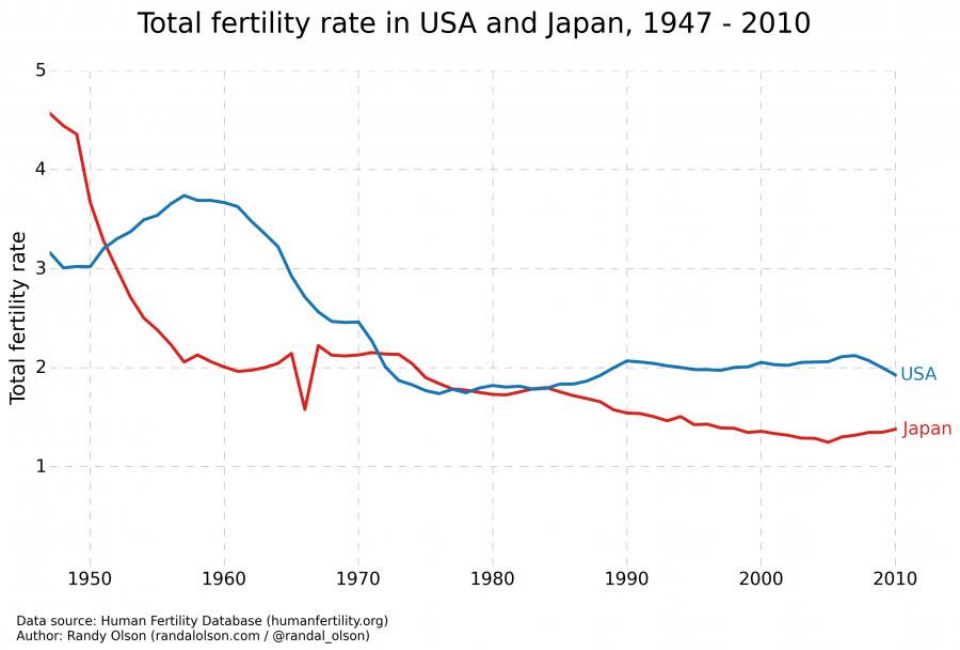
A lot of charts

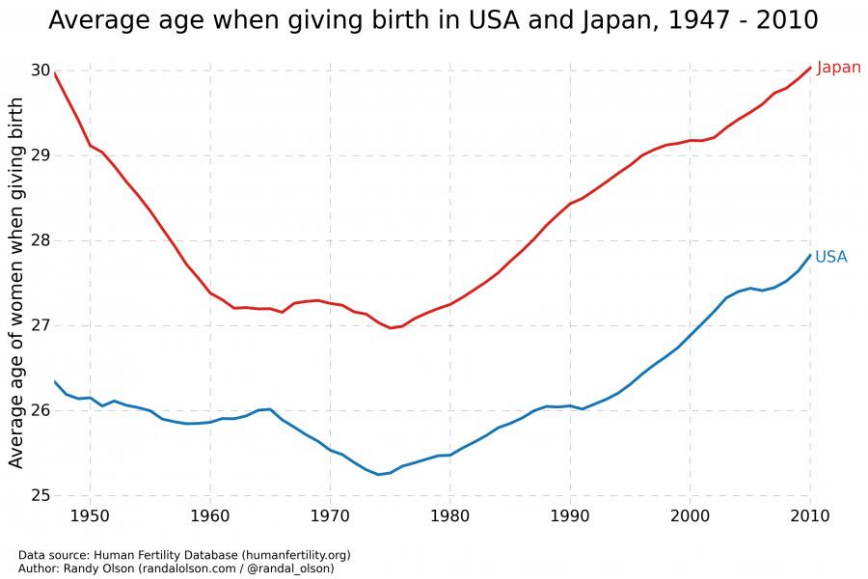
Do we need all of them?

EYES BEAT MEMORY: SMALL MULTIPLES

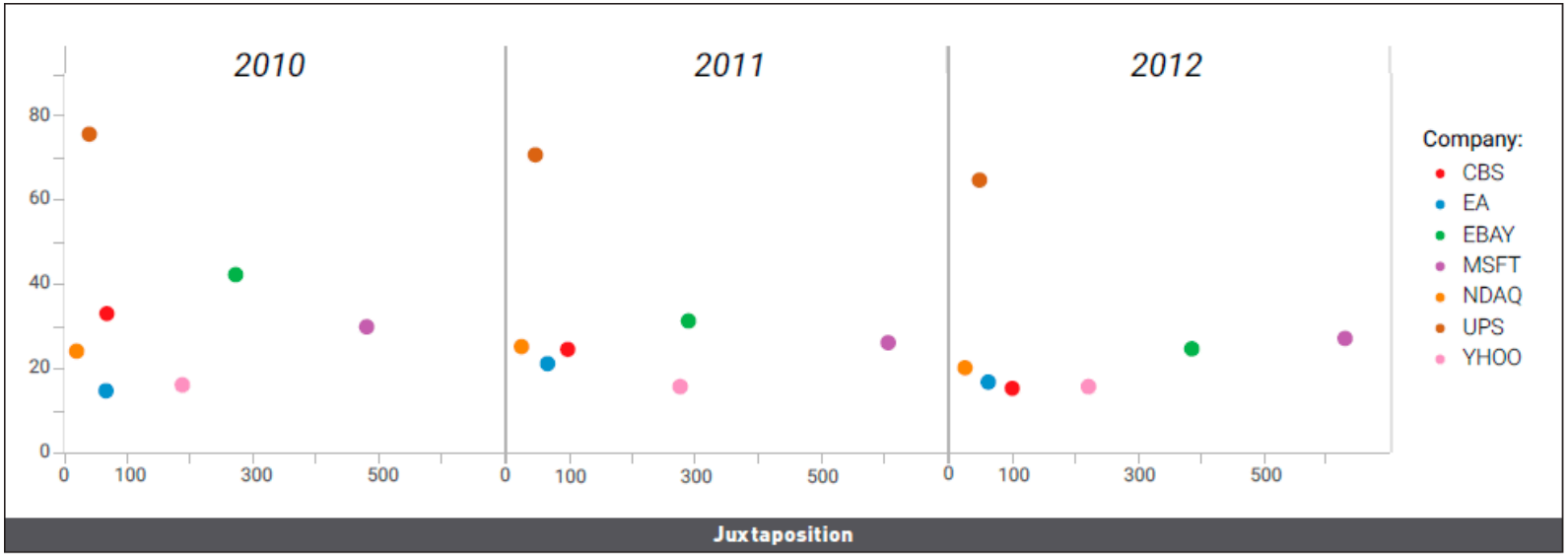


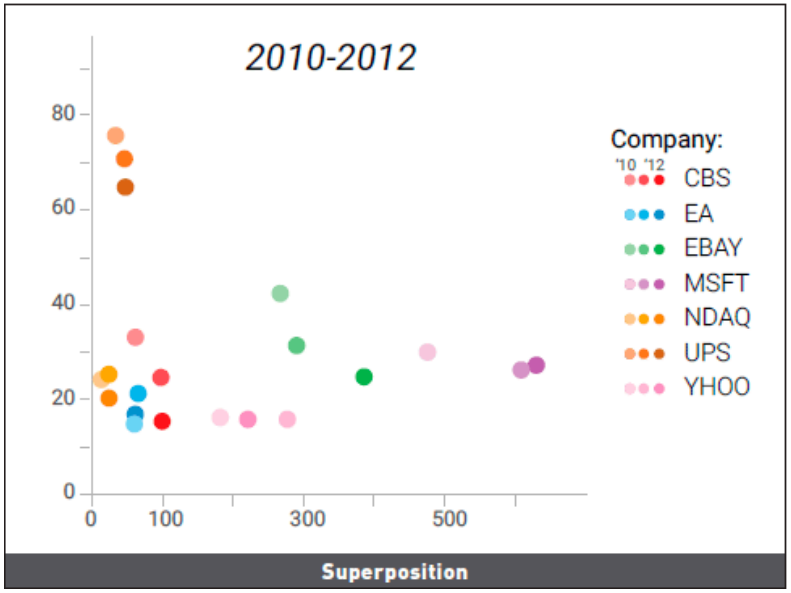
SIMPLIFY

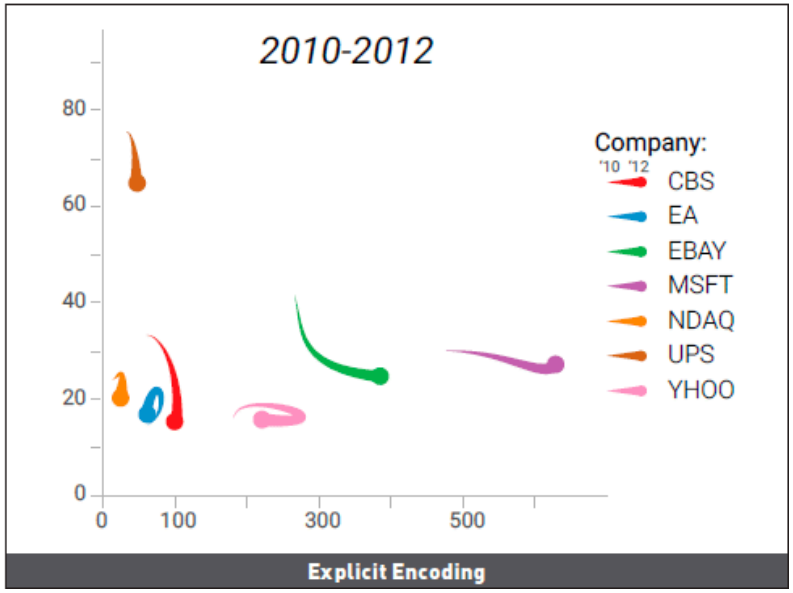

Total fertility rate in USA and Japan, 1947 - 2010
Total fertility rate
5
4
3
2
1
1950
1960
1970
1980
1990
2000
2010
USA
Japan
Data source: Human Fertility Database (humanfertility.org)
Author: Randy Olson (randalolson.com / @randal_olson)


Average age when giving birth in USA and Japan, 1947 - 2010
Average age of women when giving birth
30
29
28
27
26
25
1950
1960
1970
1980
1990
2000
2010
Japan
USA
Data source: Human Fertility Database (humanfertility.org)
Author: Randy Olson (randalolson.com / @randal_olson)

SMALL MULTIPLE DESIGN ALTERNATIVES







http://interactions.acm.org/archive/view/july-august-2018/the-good-the-bad-and-the-biased