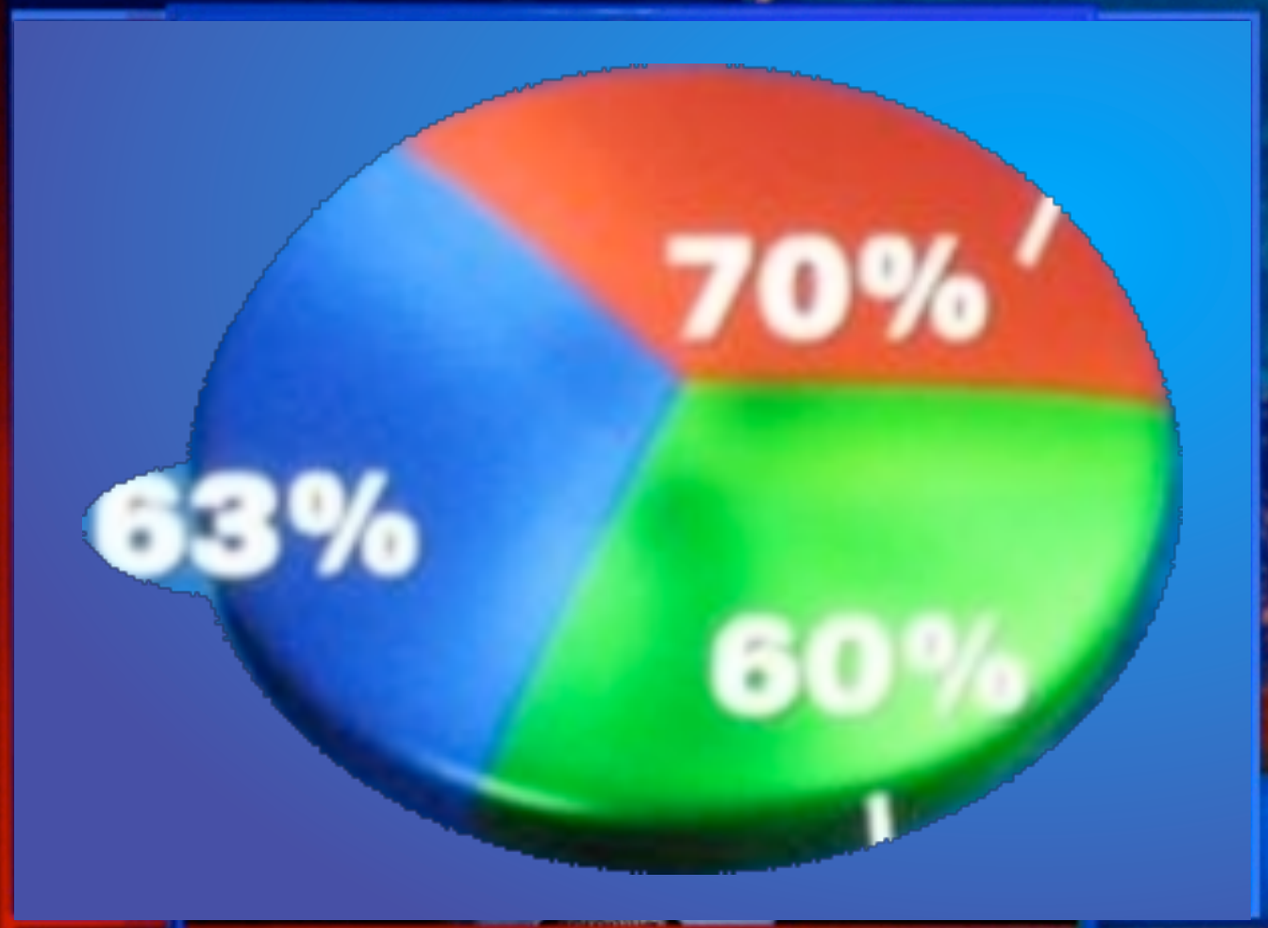
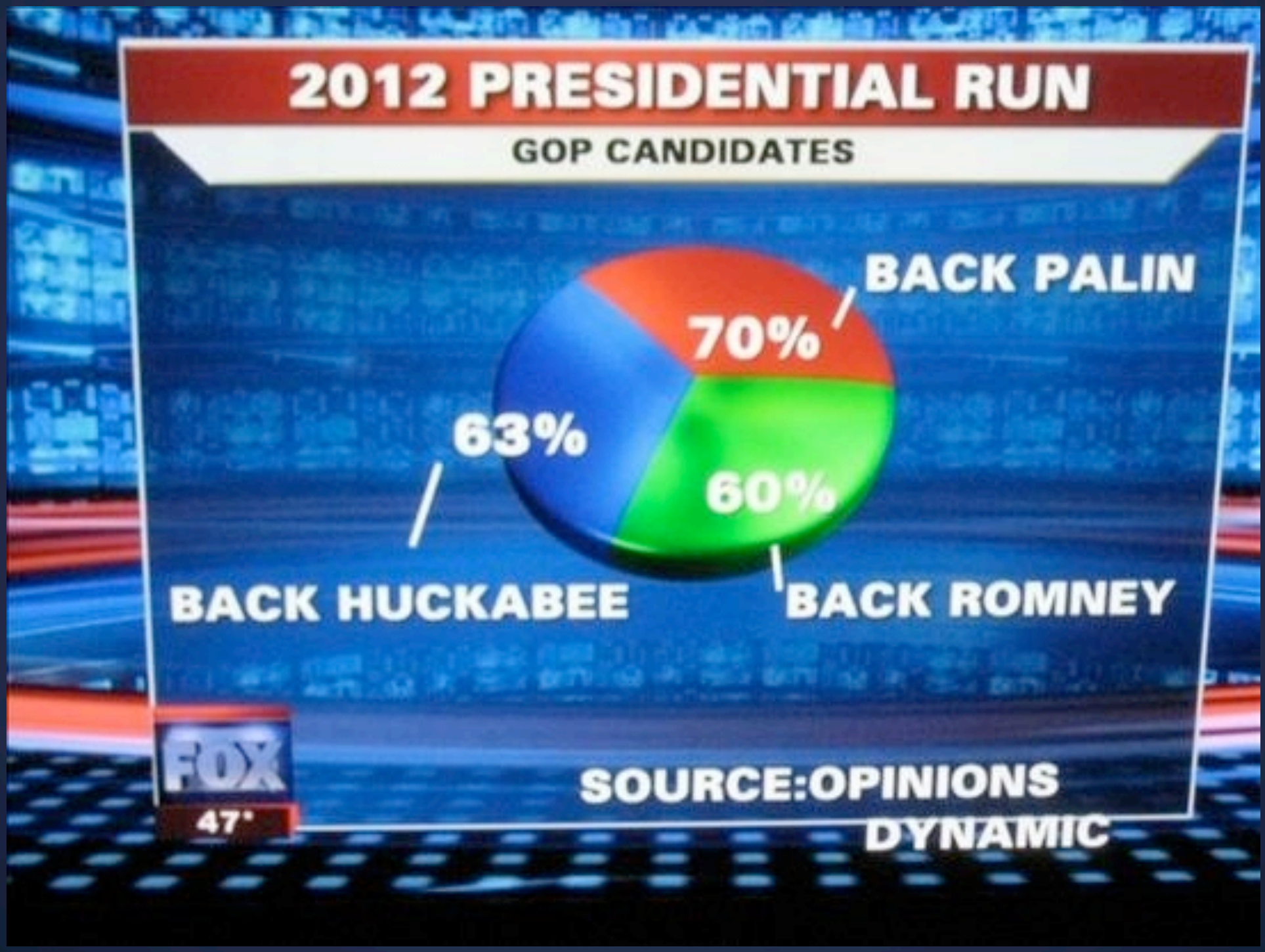
Effective Visualizations

Of Ethics and Junk

March 19, 2013 – Michael Porath (@poezn)


70%
63%
60%

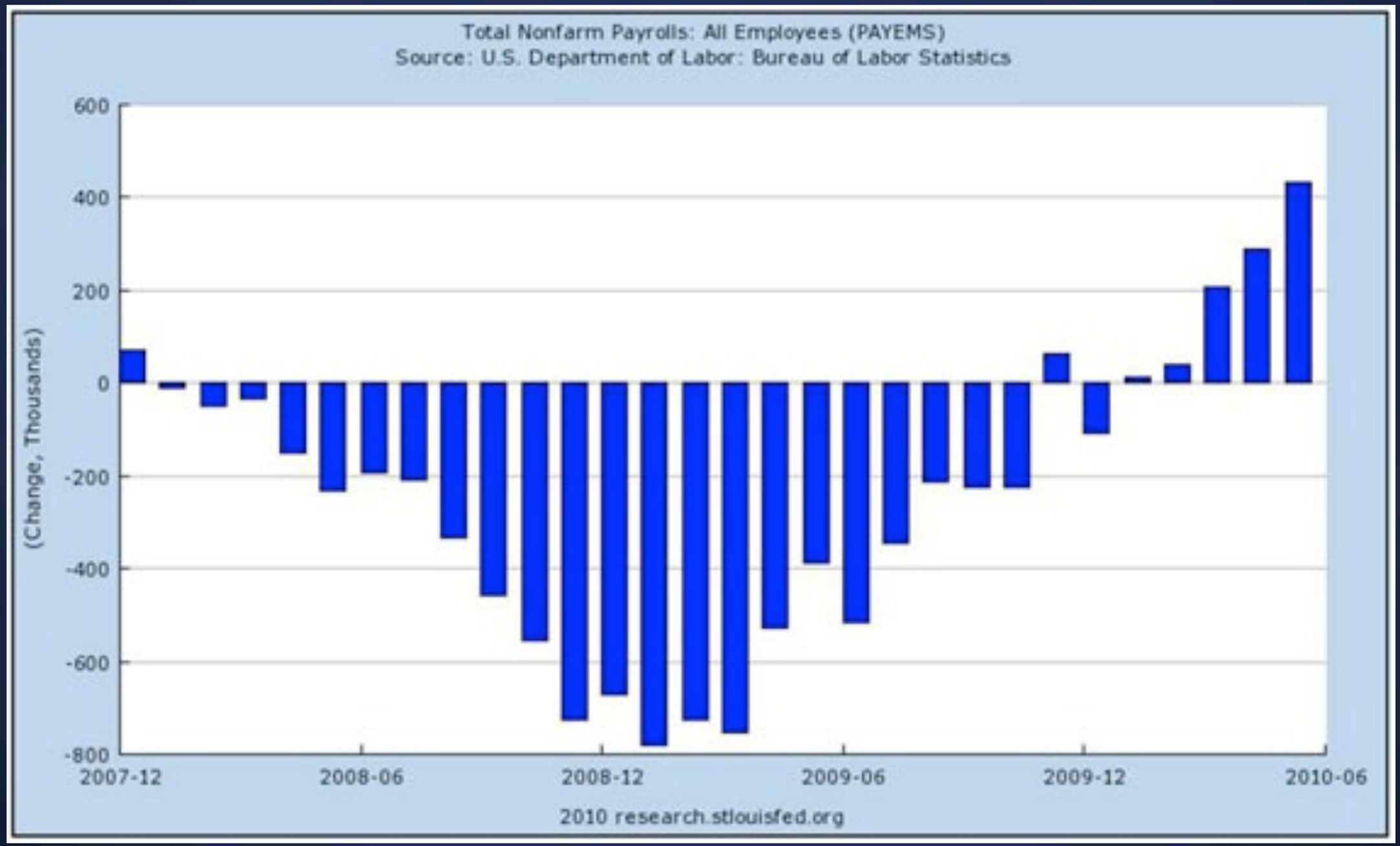
63 + 60 + 70 = 100?


2012 PRESIDENTIAL RUN
GOP CANDIDATES
BACK PALIN
70%
63%
60%
BACK HUCKABEE
BACK ROMNEY
FOX
47'
SOURCE:OPINIONS
DYNAMIC

By Quarter?

This and following slides adapted from http://mediamatters.org/blog/201006280019

Actual Job Loss by quarter

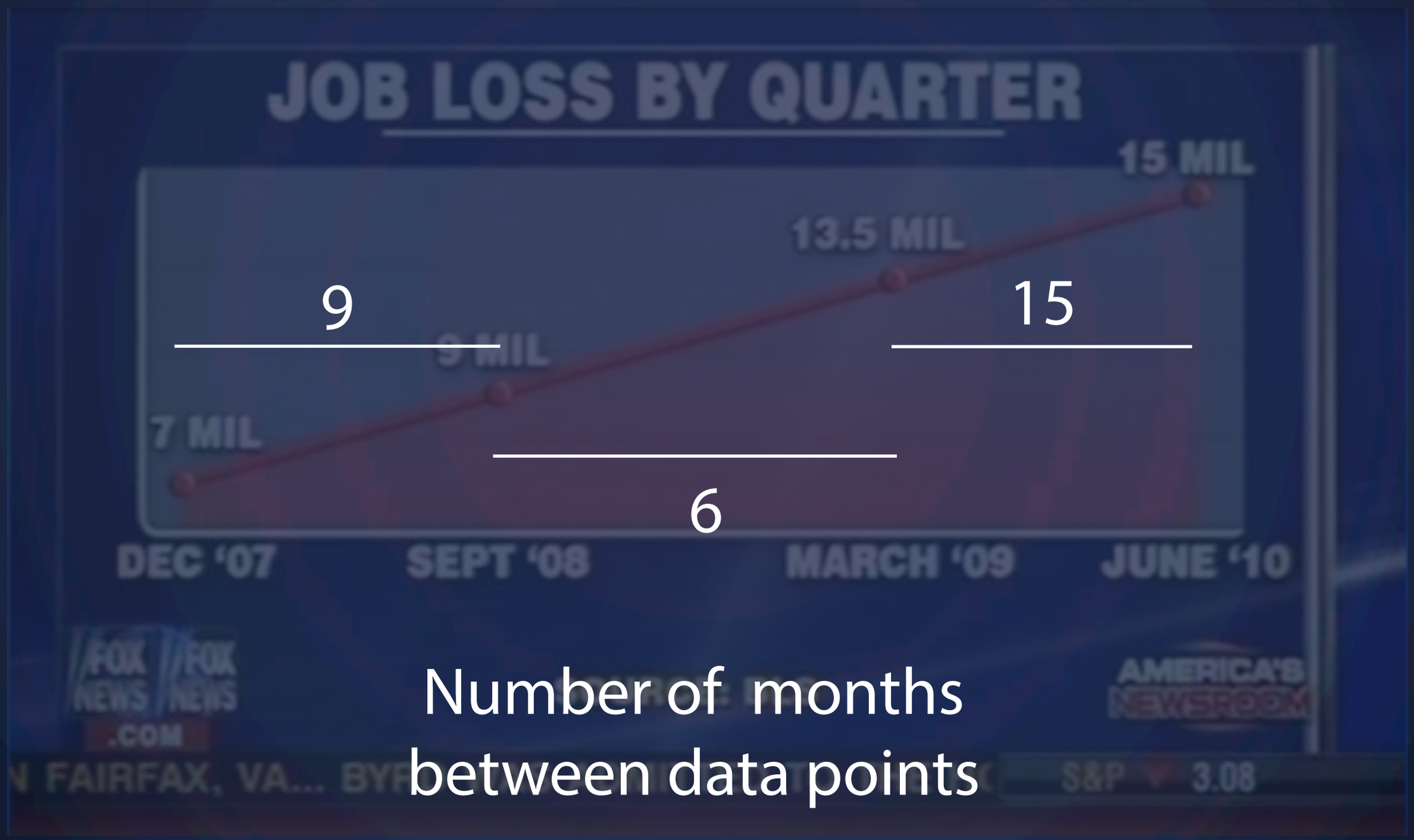


What else?

Unemployment

This and following slides adapted from http://mediamatters.org/blog/201006280019

Spacing between data points



9

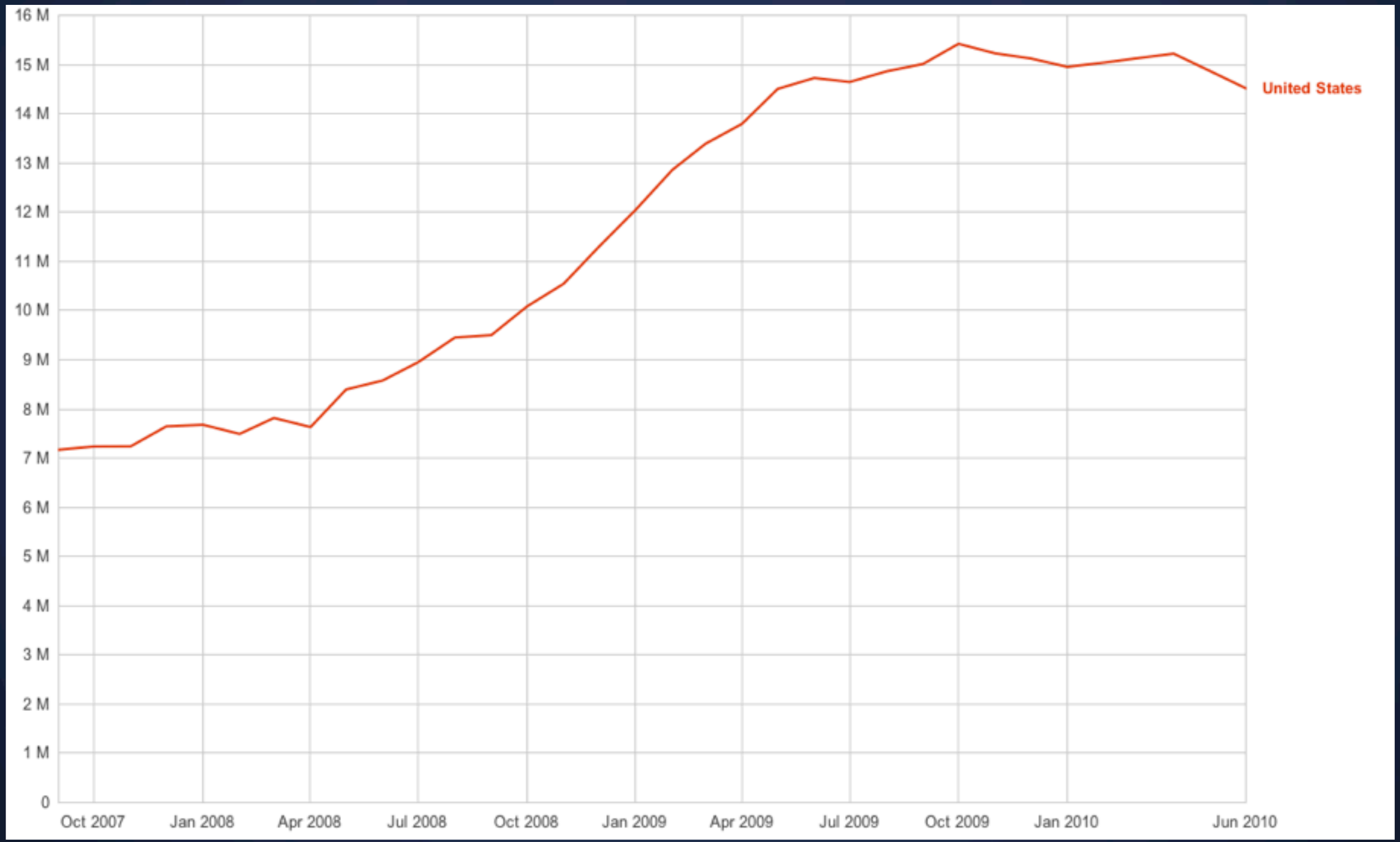
15

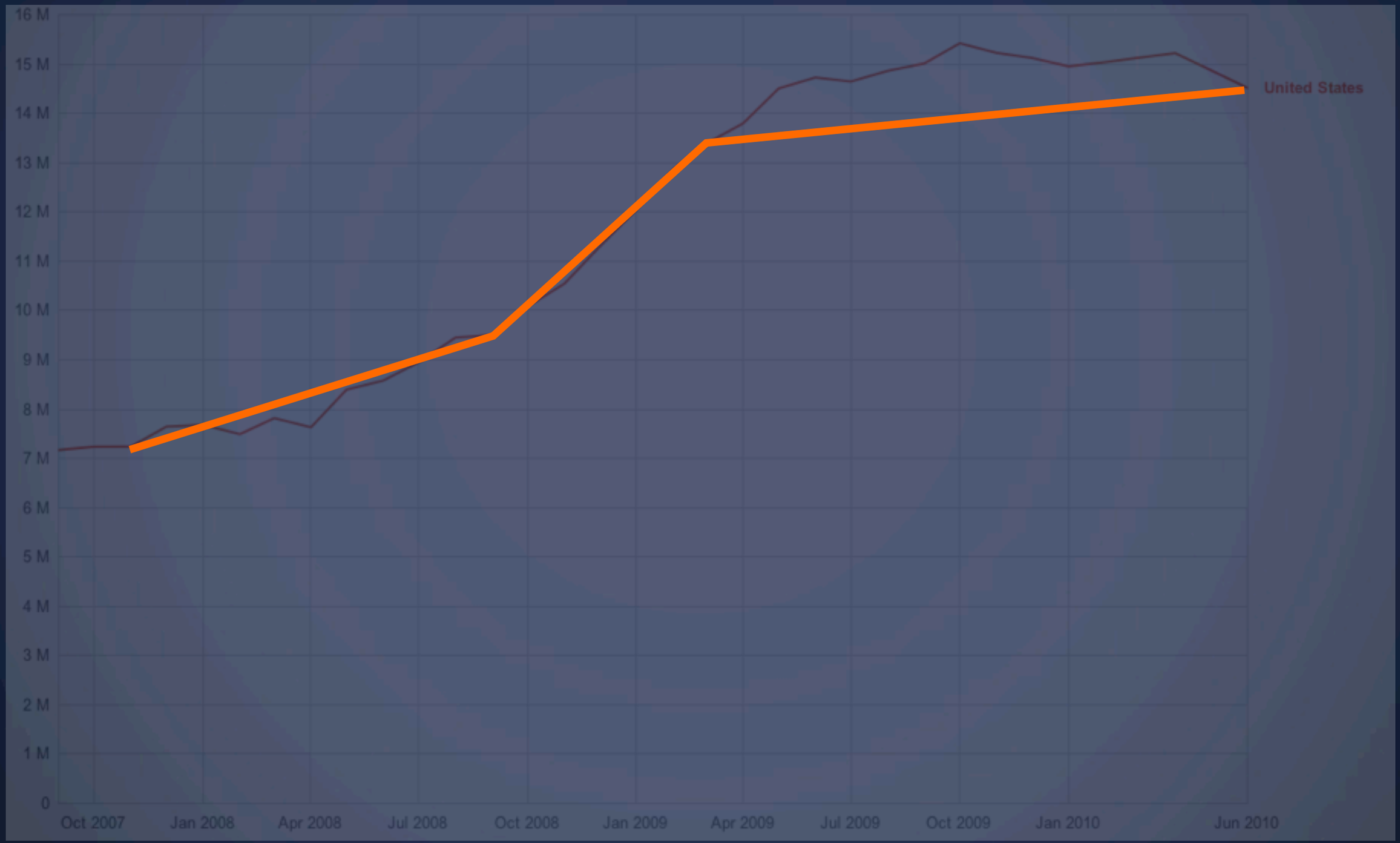
6

Number of months between data points

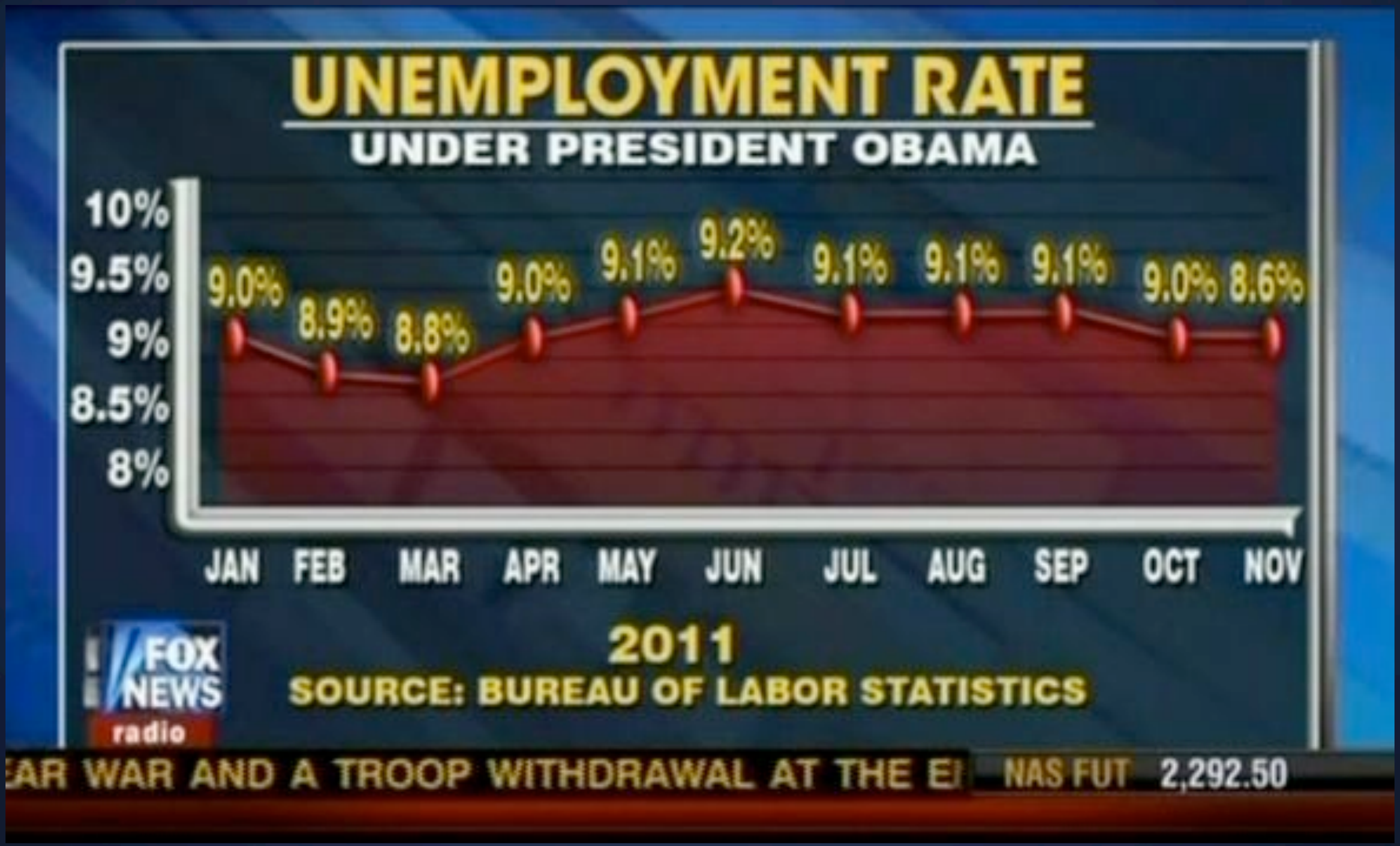
Actual Unemployment

Seasonally adjusted

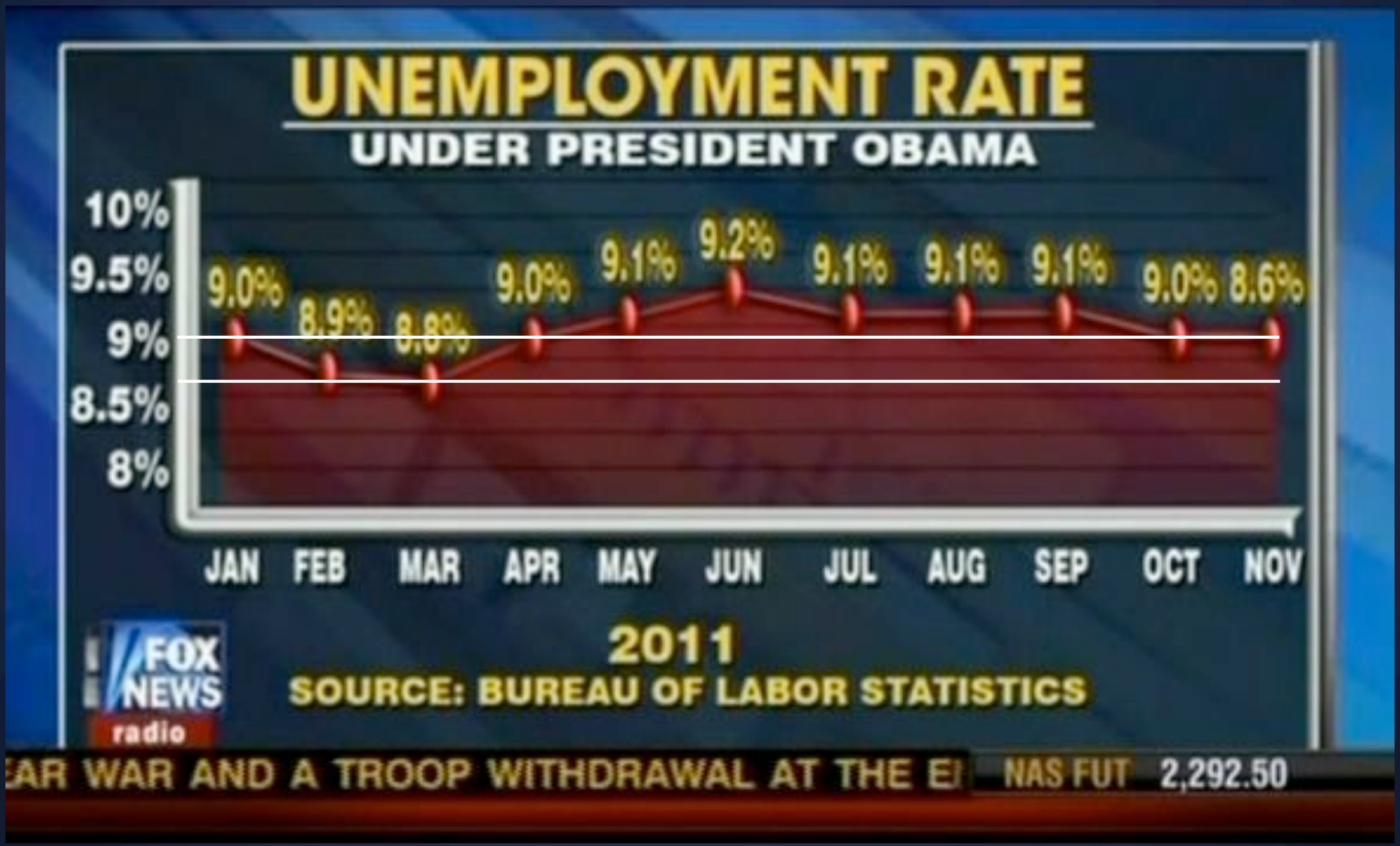



16 M
15 M
14 M
13 M
12 M
11 M
10 M
9 M
8 M
7 M
6 M
5 M
4 M
3 M
2 M
1 M
0
Oct 2007
Jan 2008
Apr 2008
Jul 2008
Oct 2008
Jan 2009
Apr 2009
Jul 2009
Oct 2009
Jan 2010
Jun 2010
United States

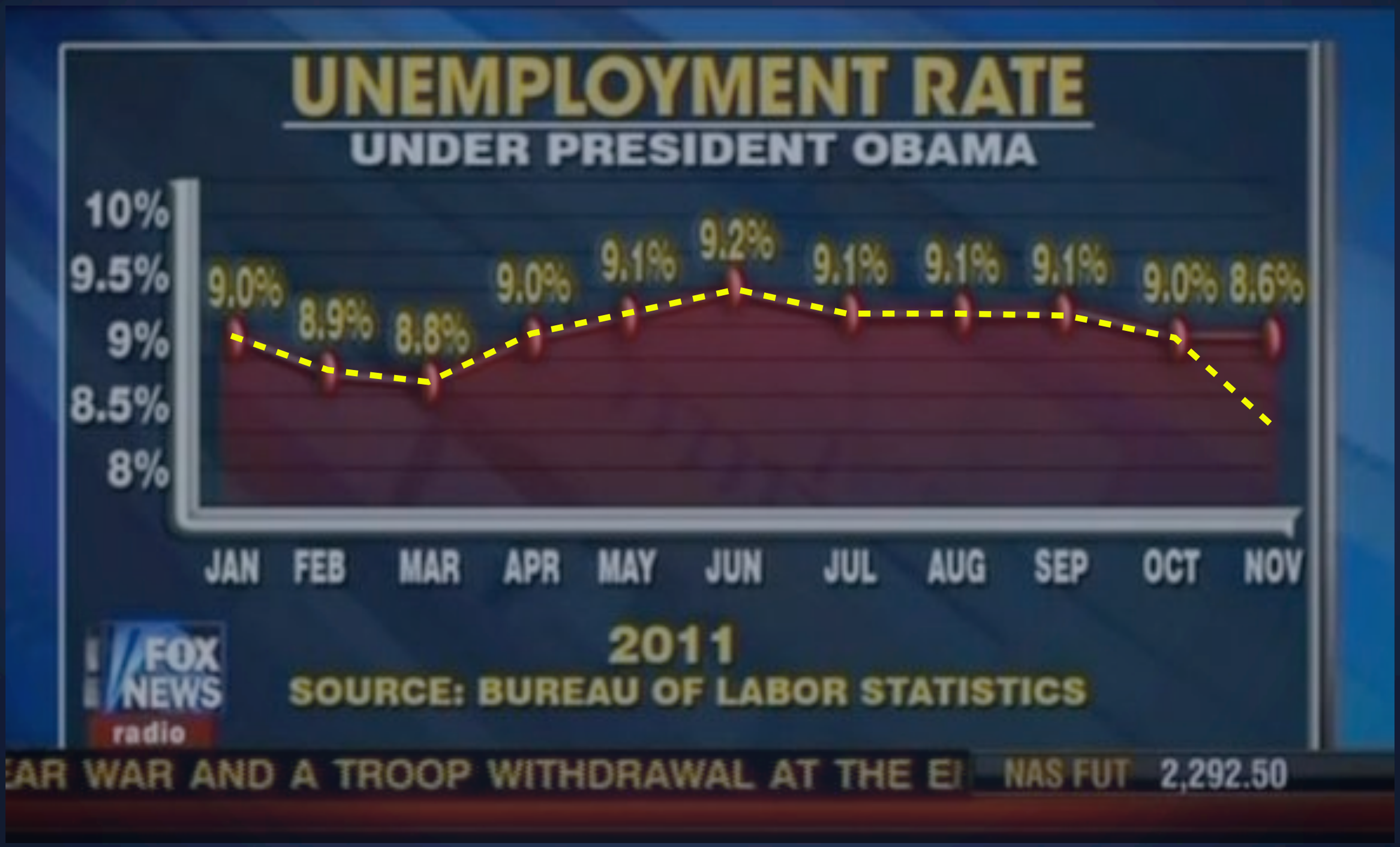
One more


UNEMPLOYMENT RATE
UNDER PRESIDENT OBAMA
10%
9.5%
9%
8.5%
8%
9.0%
8.9%
8.8%
9.0%
9.1%
9.2%
9.1%
9.1%
9.1%
9.0%
8.6%
JAN
FEB
MAR
APR
MAY
JUN
JUL
AUG
SEP
OCT
NOV
2011
SOURCE: BUREAU OF LABOR STATISTICS
FOX NEWS radio
AR WAR AND A TROOP WITHDRAWAL AT THE E
NAS FUT 2,292.50

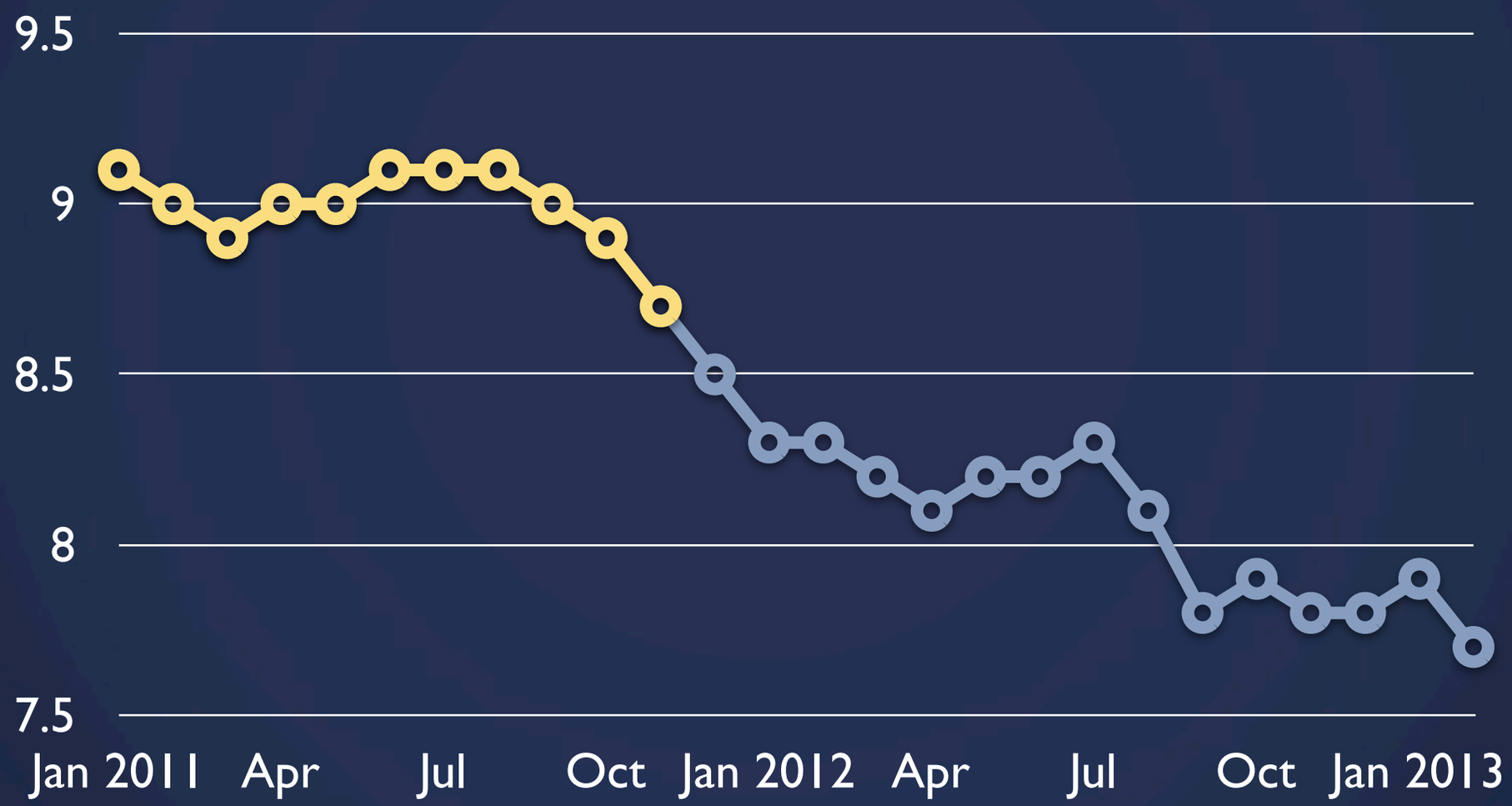
An honest mistake?



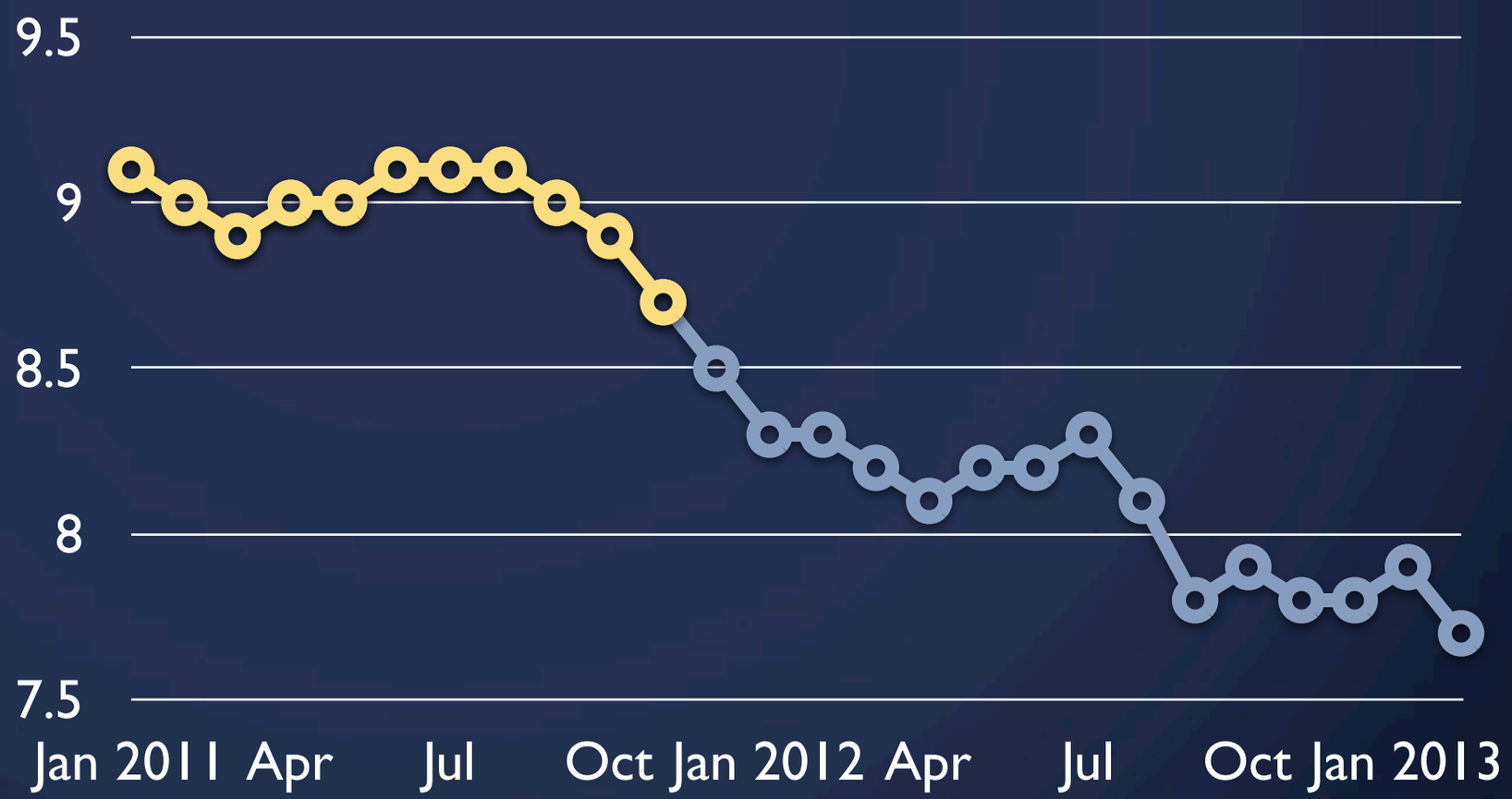
An honest mistake?



Actual unemployment rate



Actual unemployment rate



Actual unemployment rate

Doesn't that look a lot different?

No need to lie

You only have to highlight. (NY Times, October 2012)

How a **Democrat** Might See Things

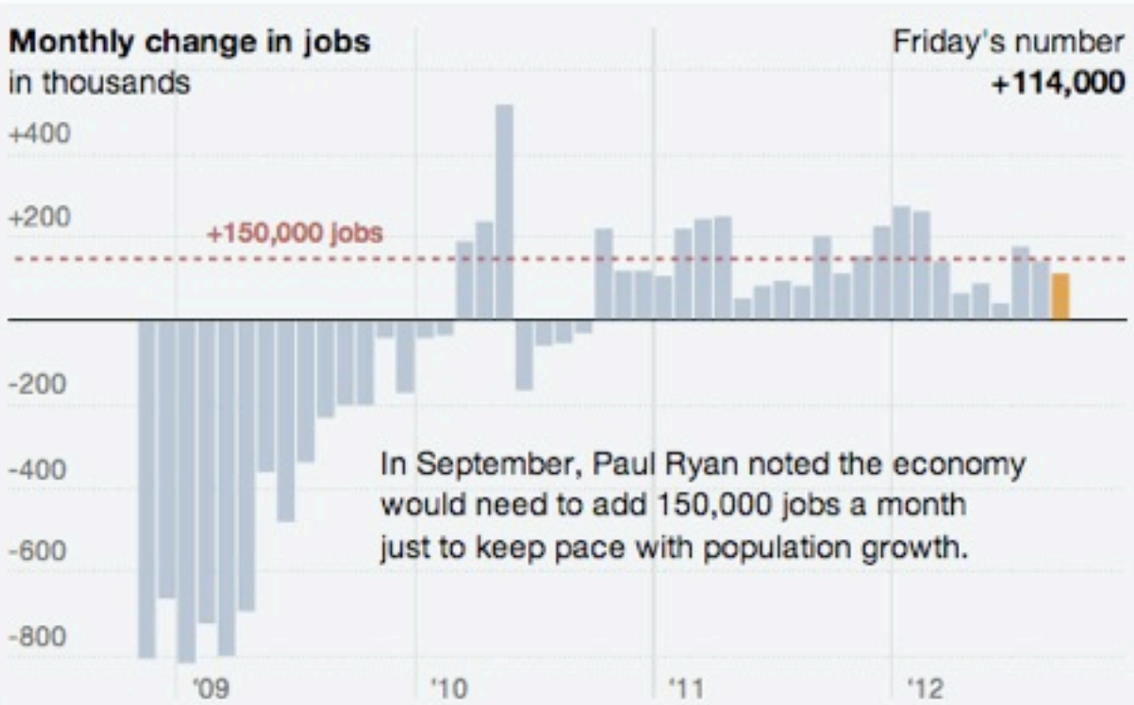
September Jobs Report

How a **Republican** Might See Things

Job growth has been weak.

Monthly change in jobs
in thousands

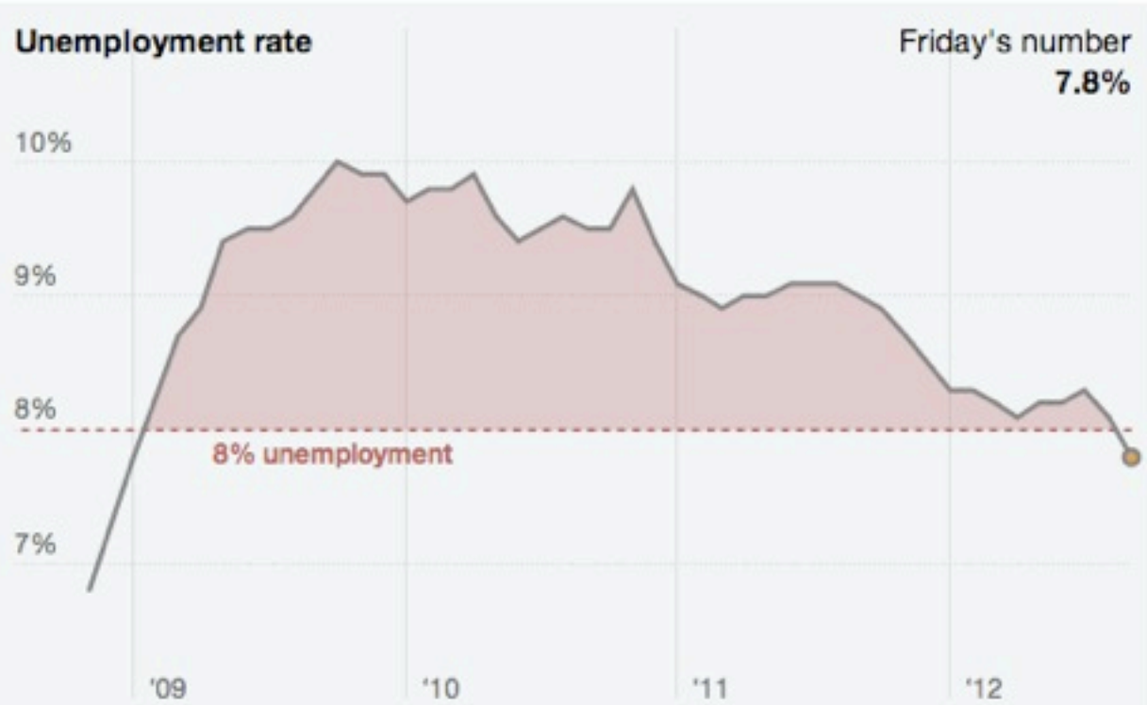
Friday's number
+114,000



The rate was above 8 percent for 43 months.

Unemployment rate

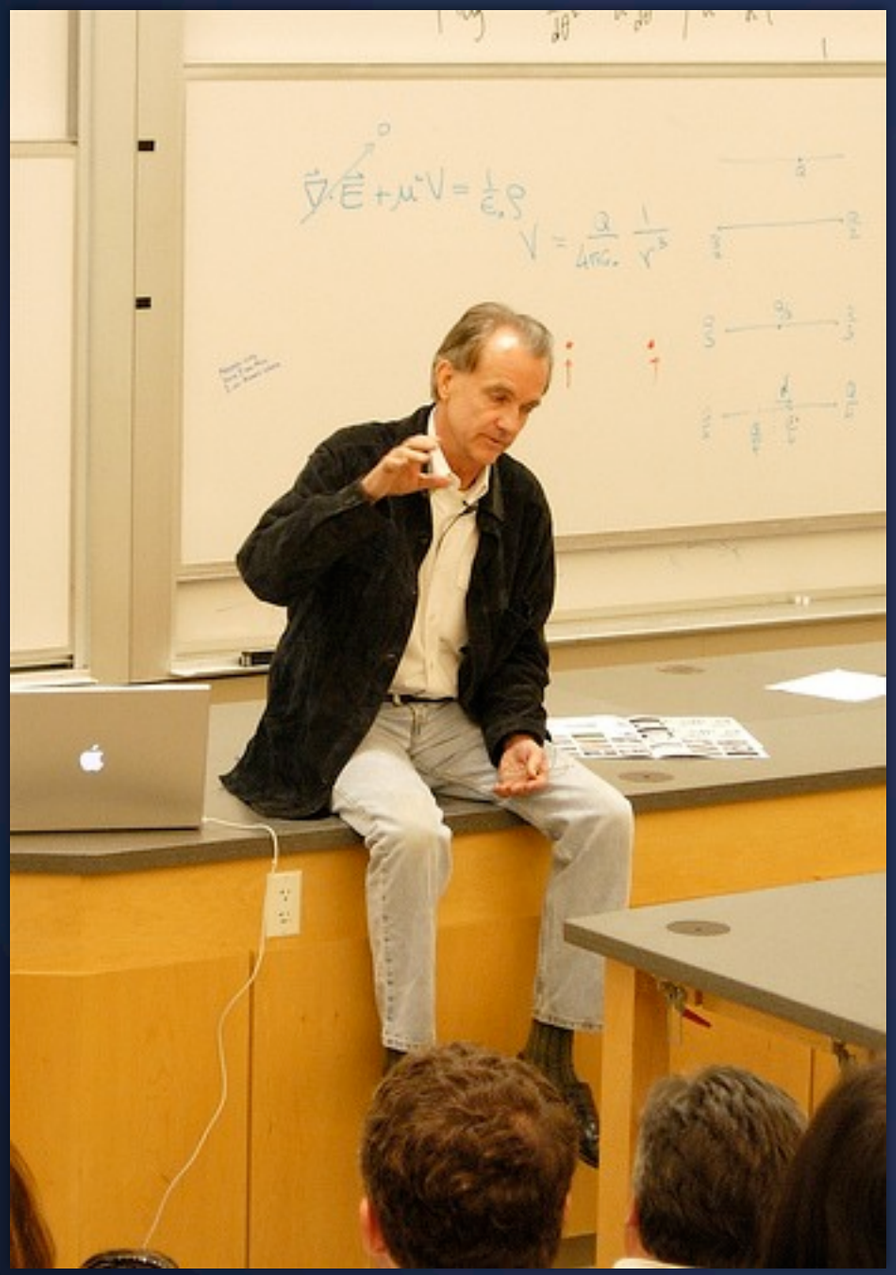
Friday's number
7.8%



http://www.nytimes.com/interactive/2012/10/05/business/economy/one-report-diverging-perspectives.html?smid=tw-share&_r=0

Edward Tufte

A bit outdated... but still true



Graphical Integrity

"Tell the truth"

Aesthetics

"Do it well"

Graphical Integrity

Graphical Integrity

According to Tufte

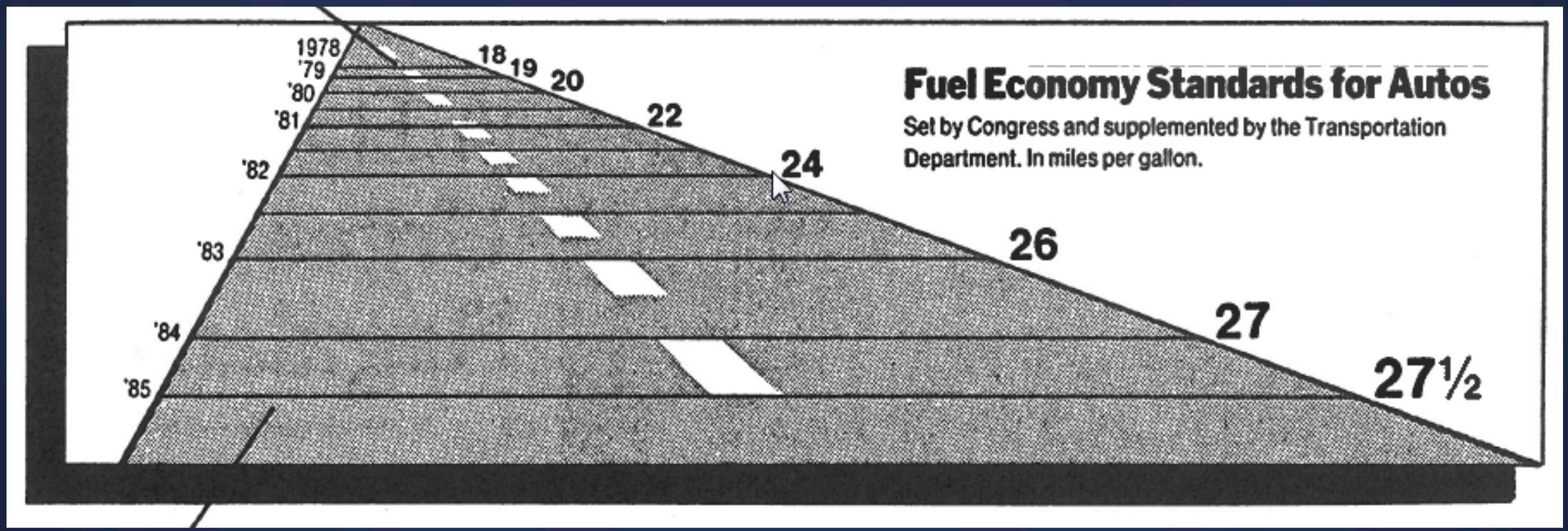
Well designed presentation of interesting data

Complex ideas communicated with clarity, precision and efficiency

Gives the viewer greatest number of ideas in the shortest amount of time, with the least ink and smallest space

Nearly always multivariate

What's wrong with this picture?



New York Times, August 1978

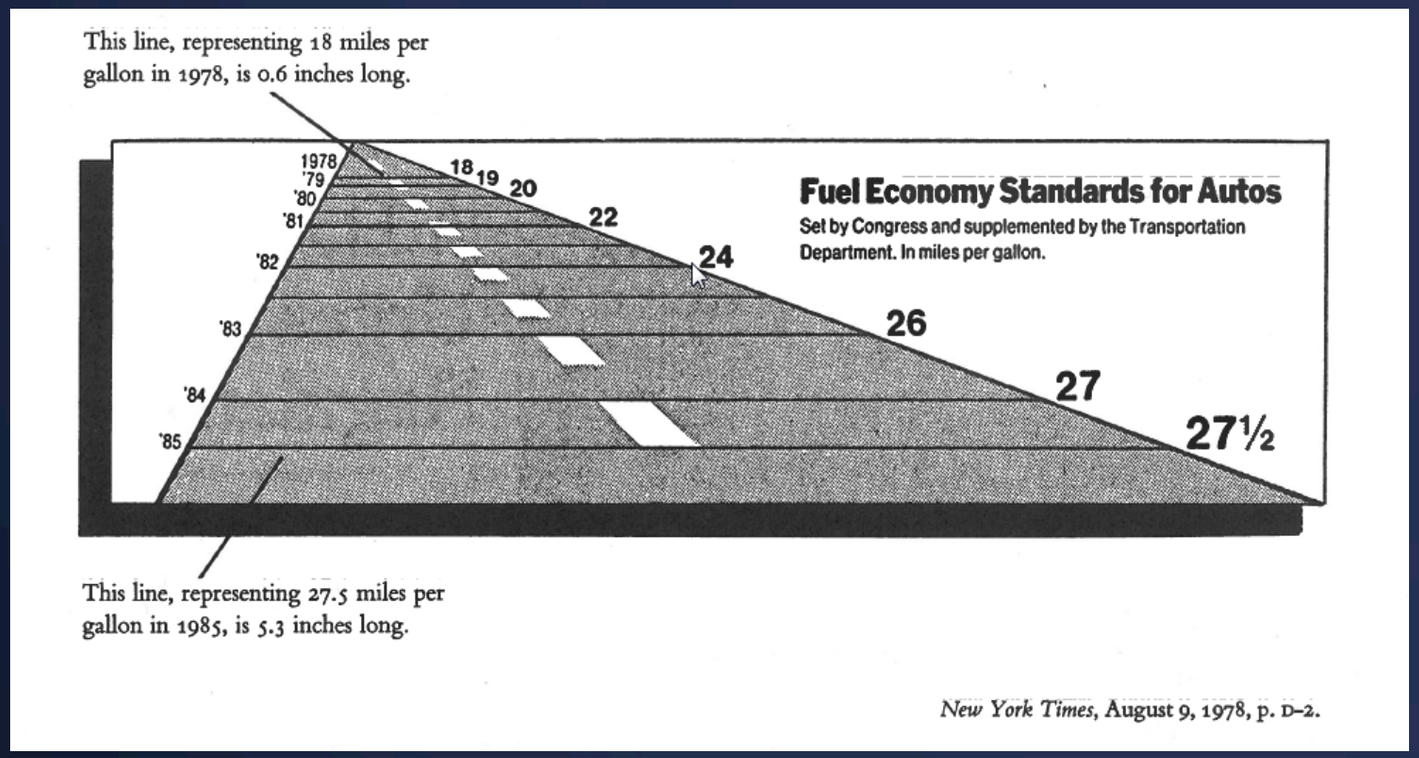
What's wrong with this picture?

Lie Factor

$$Lie\ Factor = \frac{Size\ of\ effect\ in\ graphic}{Size\ of\ effect\ in\ data}$$

$$\frac{5.3in - 0.6in}{0.6in} = 7.83$$

$$\frac{27.5 - 18.0}{18.0} = 0.53$$



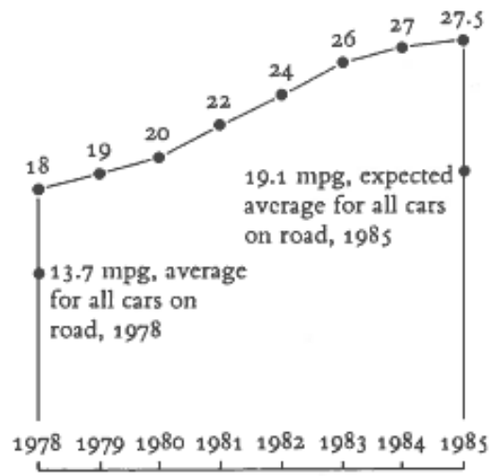
New York Times, August 9, 1978, p. D–2.

Lie factor = 14.8

Cleaned Up Version

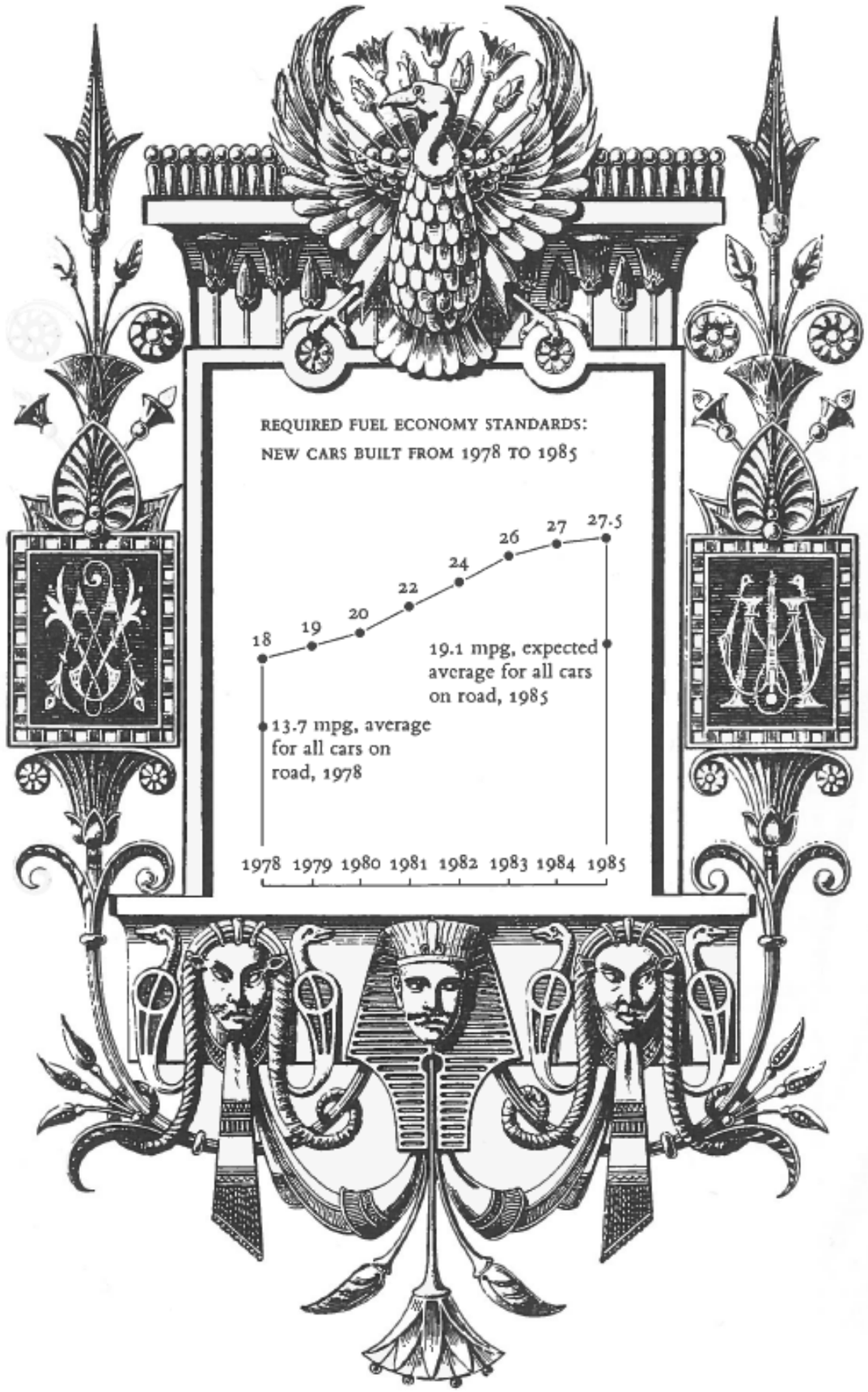
Better?

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



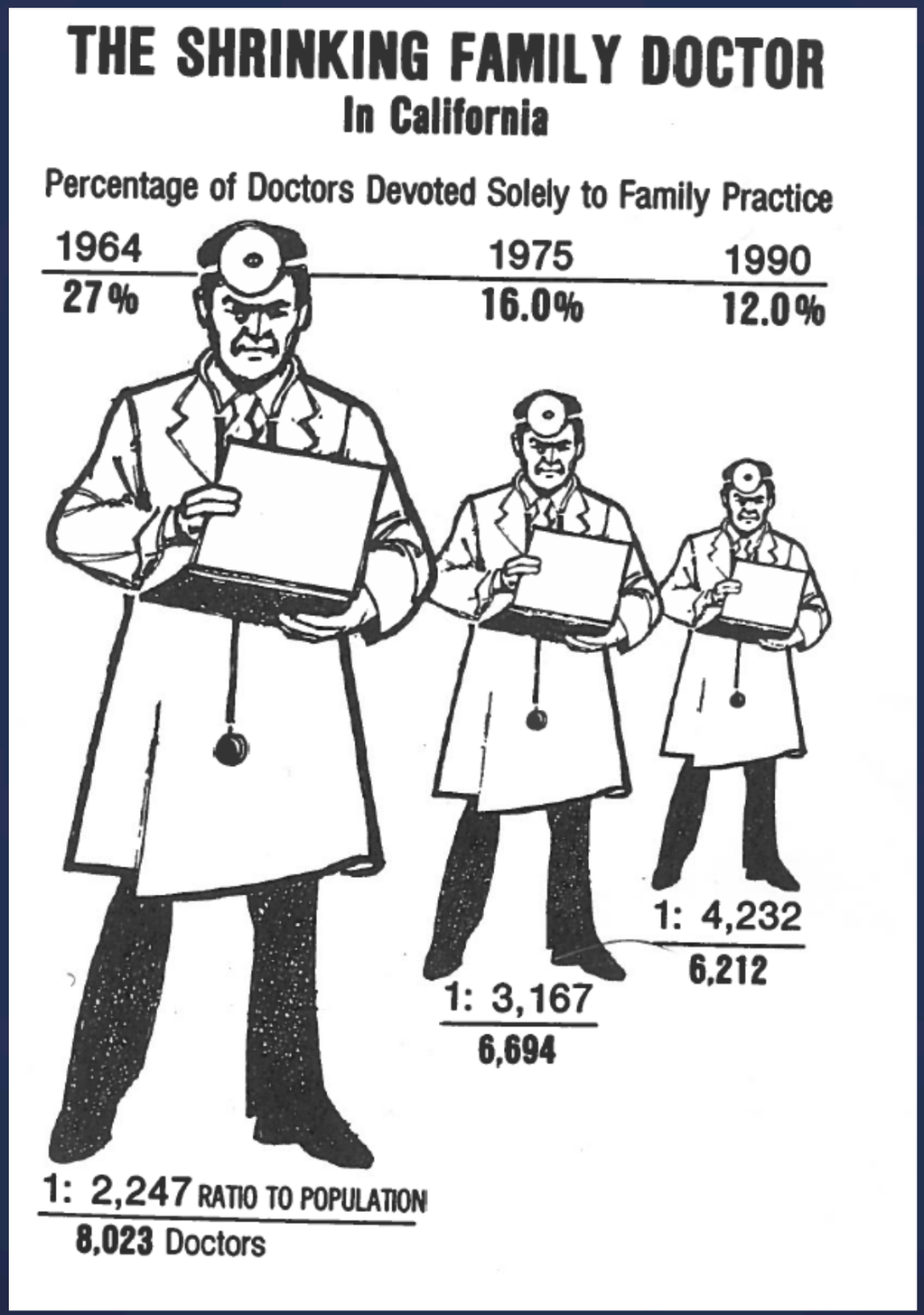
Cleaned Up Version

Better?



Scale

Remind you of something?



Scale

Remind you of something?

27%

1964

16%

1975

12%

1990

Scale

Remind you of something?

27%

1964

+69%

16%

1975

+33%

12%

1990

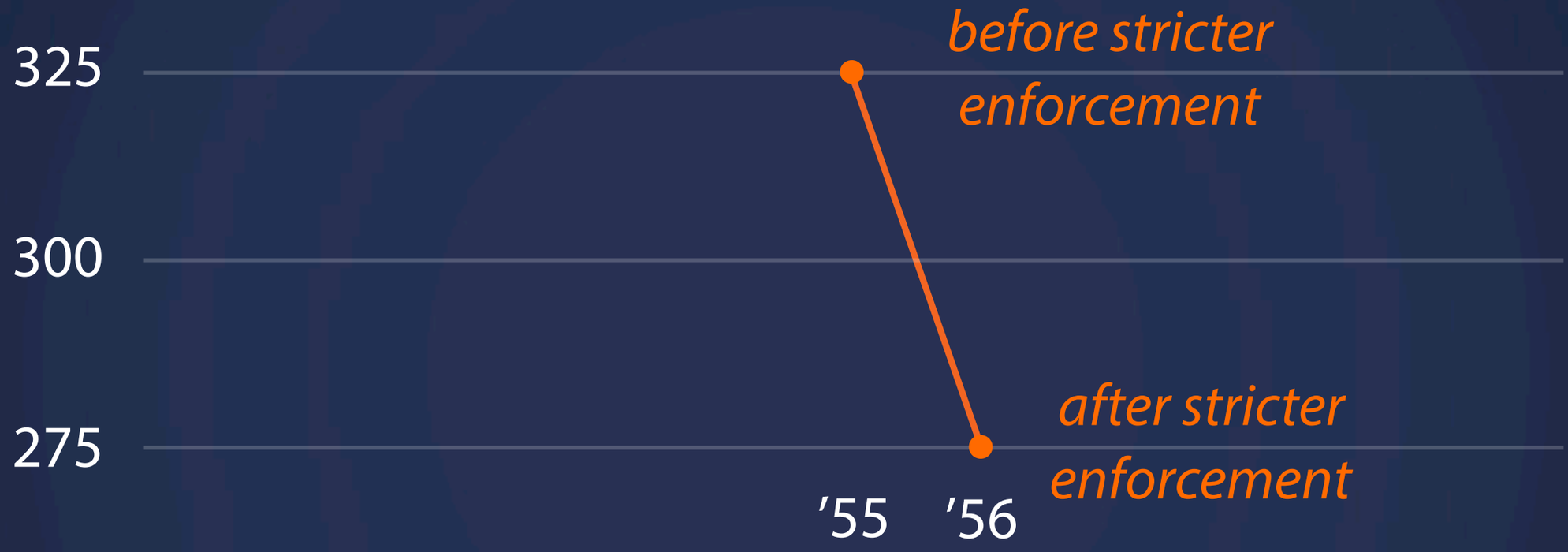
Principle 1

Don’t distort. Use the appropriate number of dimensions

(1 dimensional data = 1 dimensional graphic)

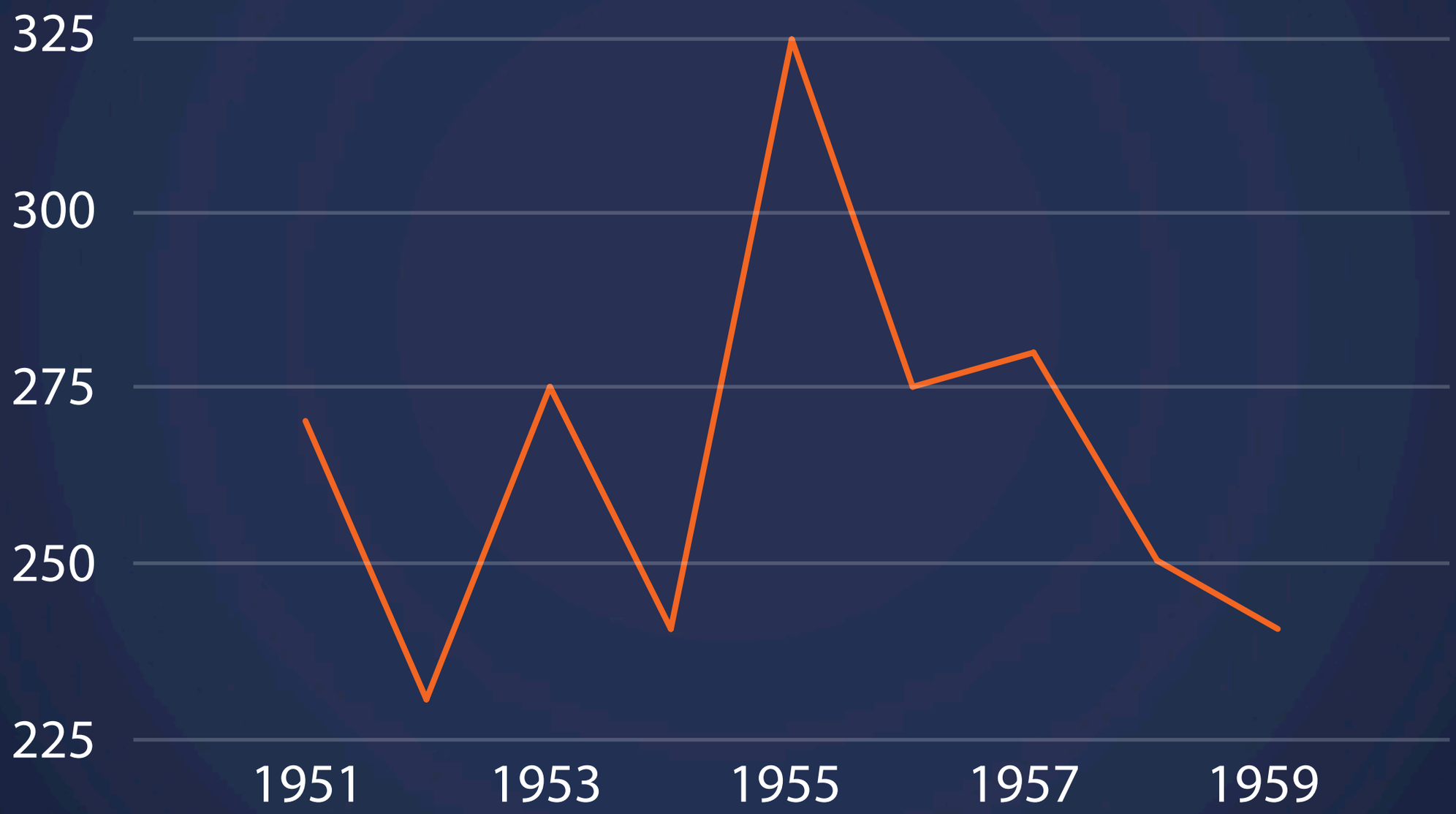
Context is key

Connecticut Traffic Deaths



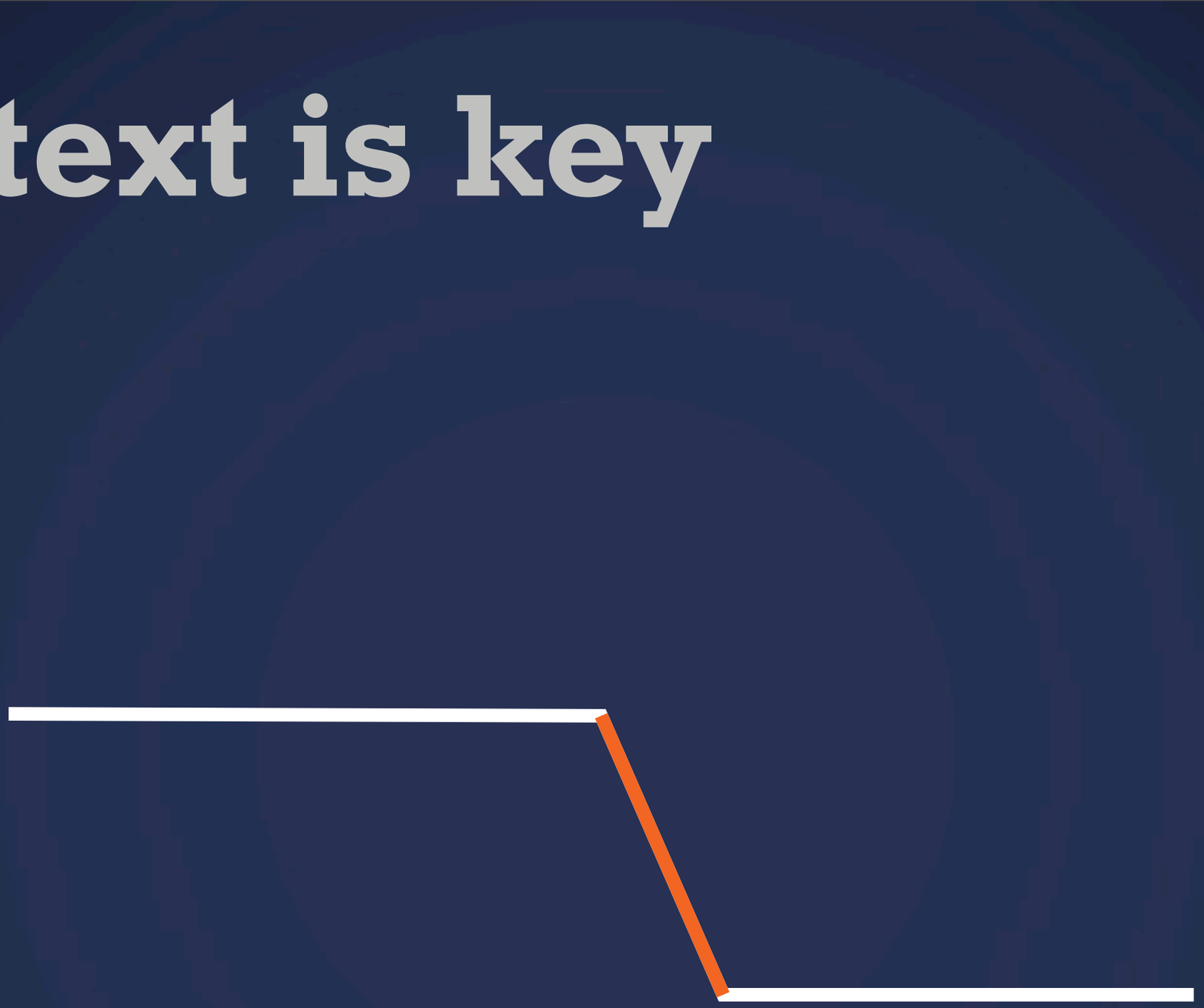
Context is key

Connecticut Traffic Deaths, 1951-1959



Context is key

Context is key



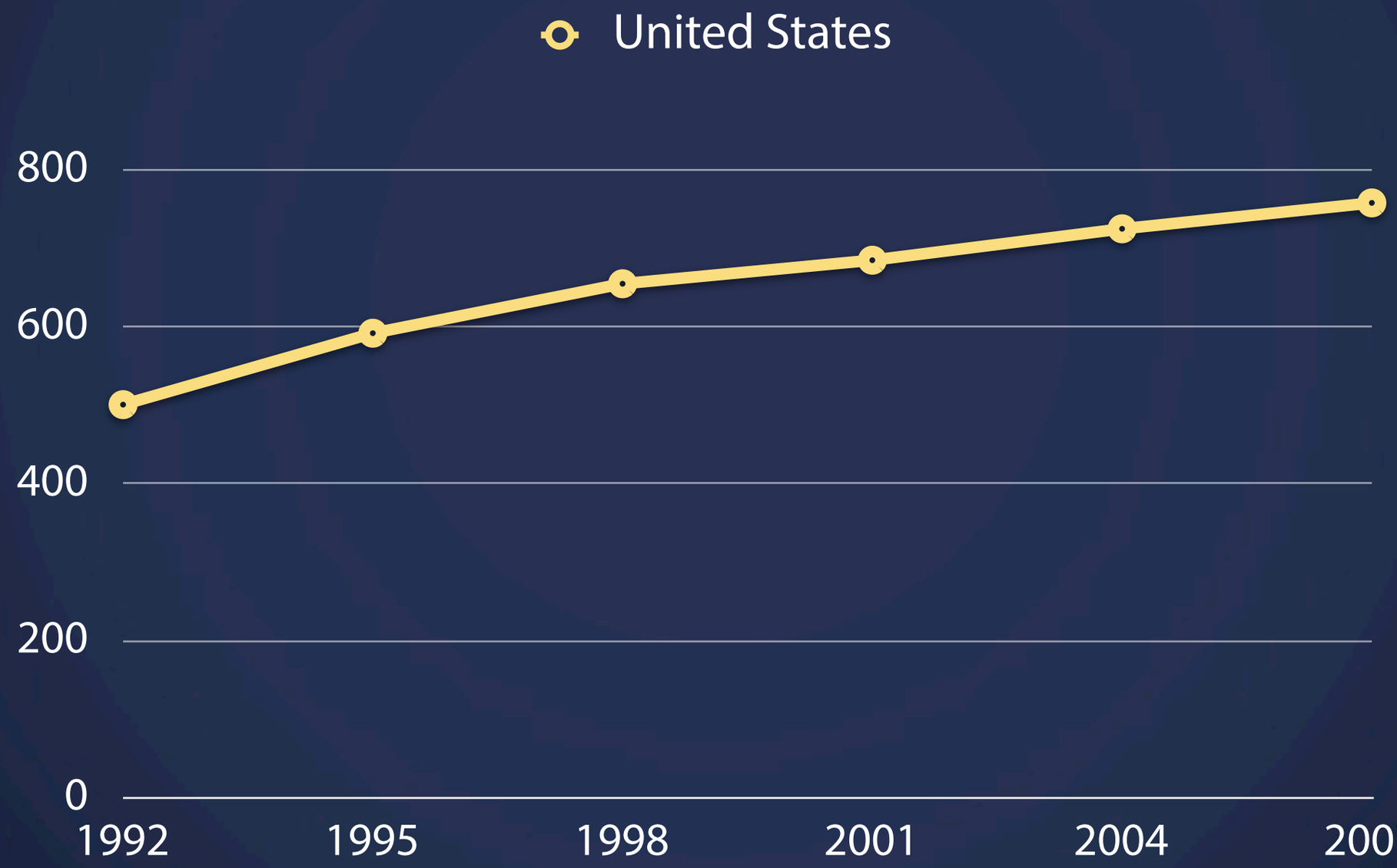
Context is key

Context is key

Context is key

Comparisons

Incarceration Rate (number of inmates per 100,000)

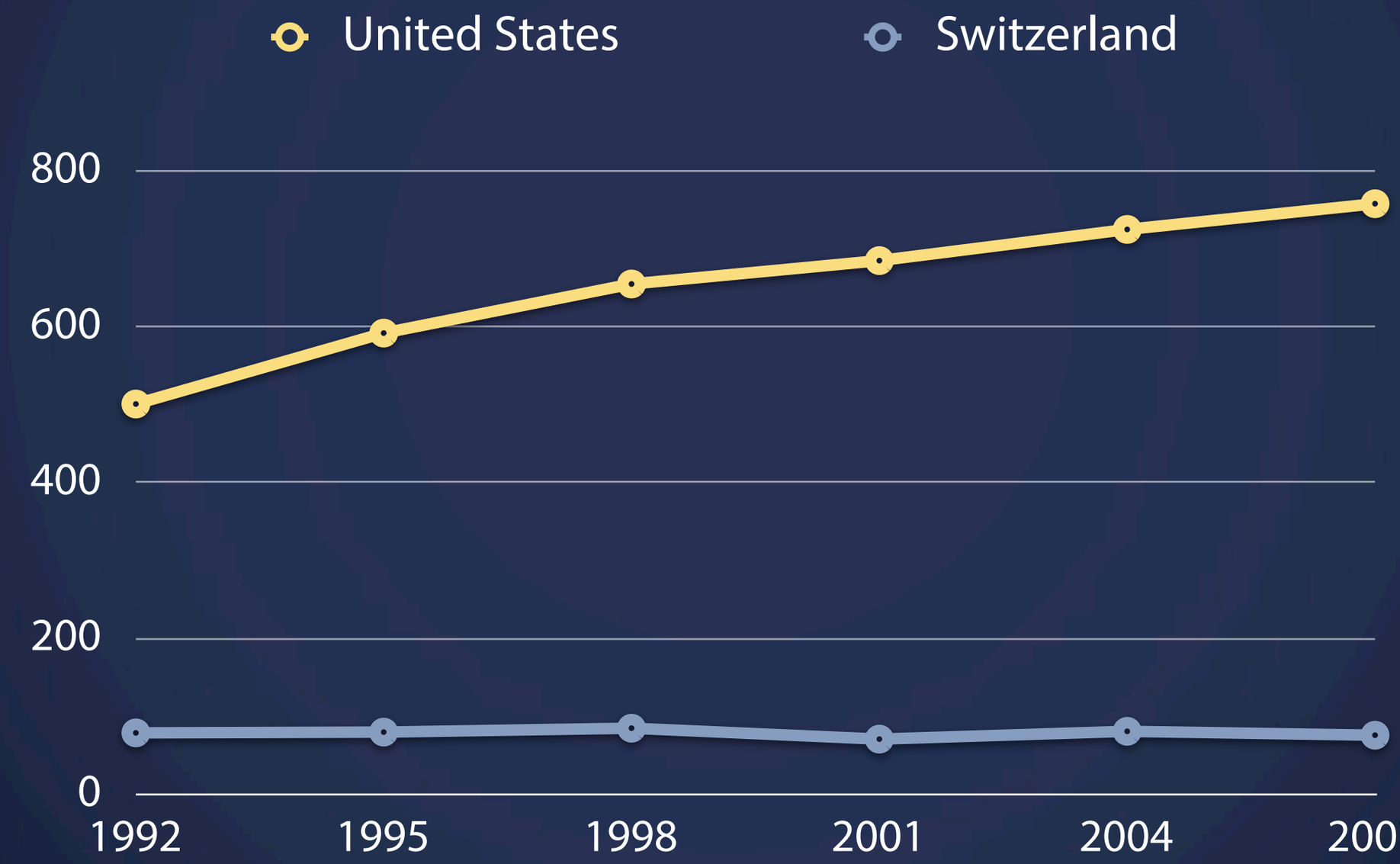


Data from http://www.prisonstudies.org/info/worldbrief/

Context is key

Comparisons

Incarceration Rate (number of inmates per 100,000)

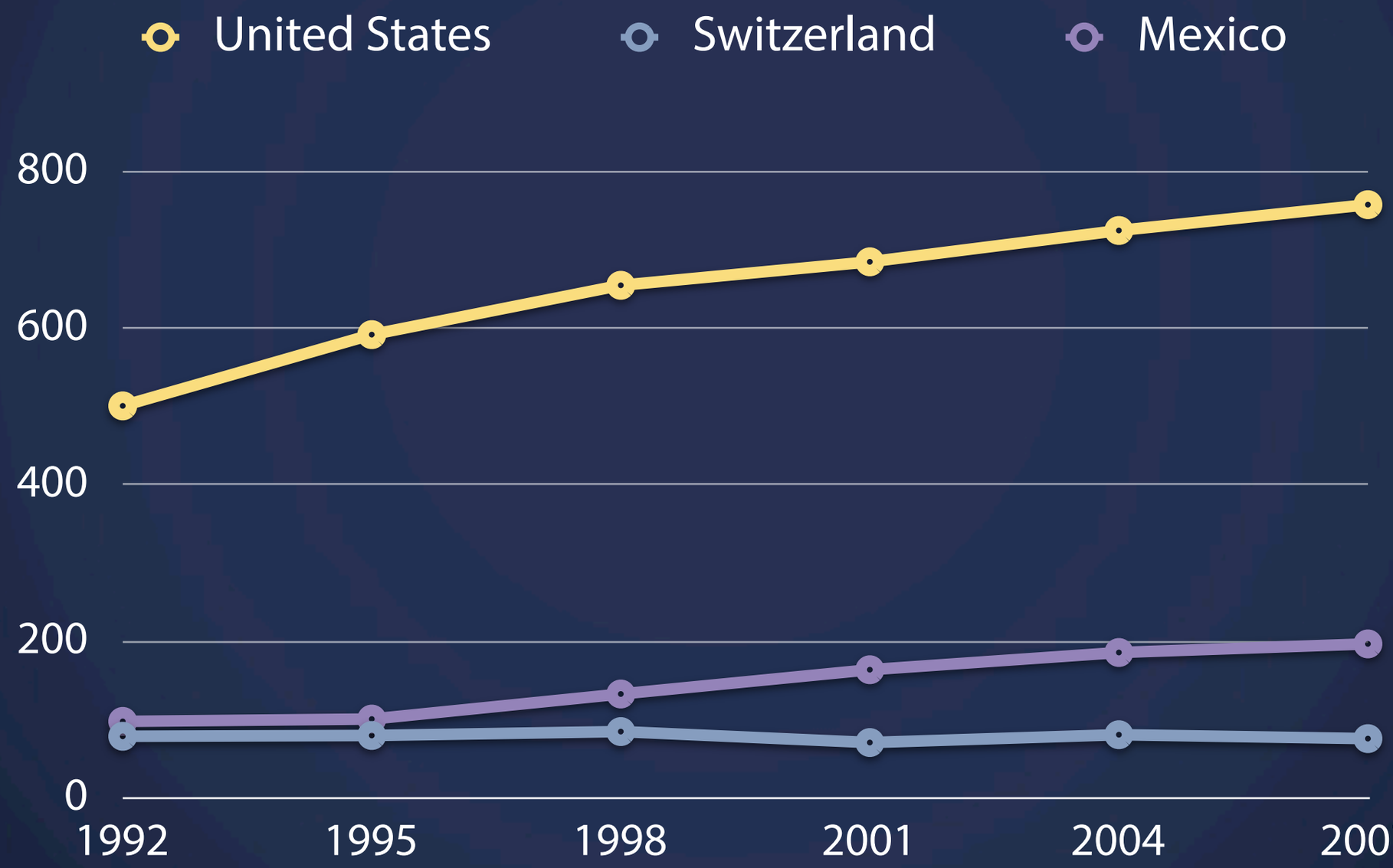


Data from http://www.prisonstudies.org/info/worldbrief/

Context is key

Comparisons

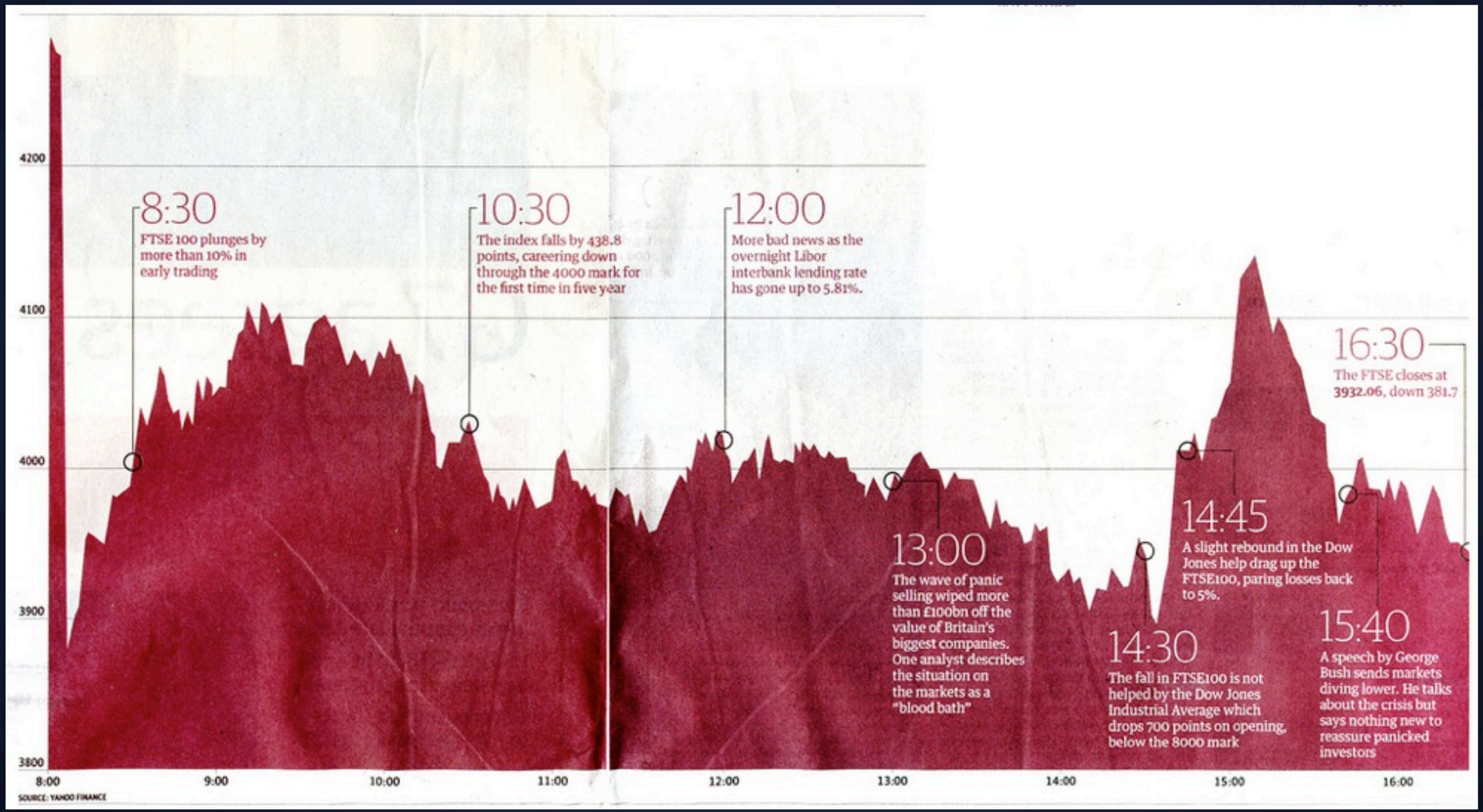
Incarceration Rate (number of inmates per 100,000)



Data from http://www.prisonstudies.org/info/worldbrief/

Context is key

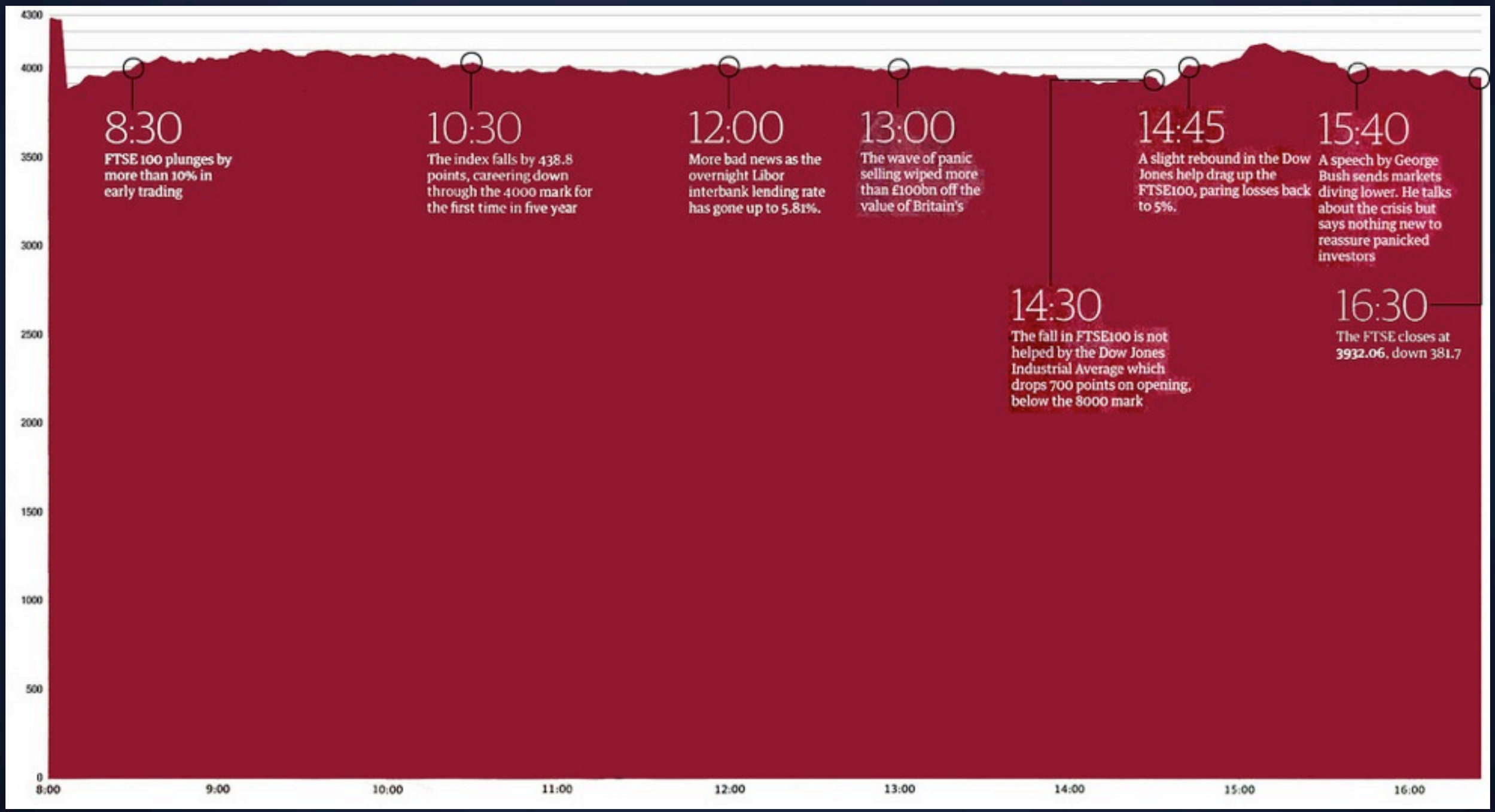
Scale



http://www.flickr.com/photos/philgyford/2932821996/in/photostream/

Context is key

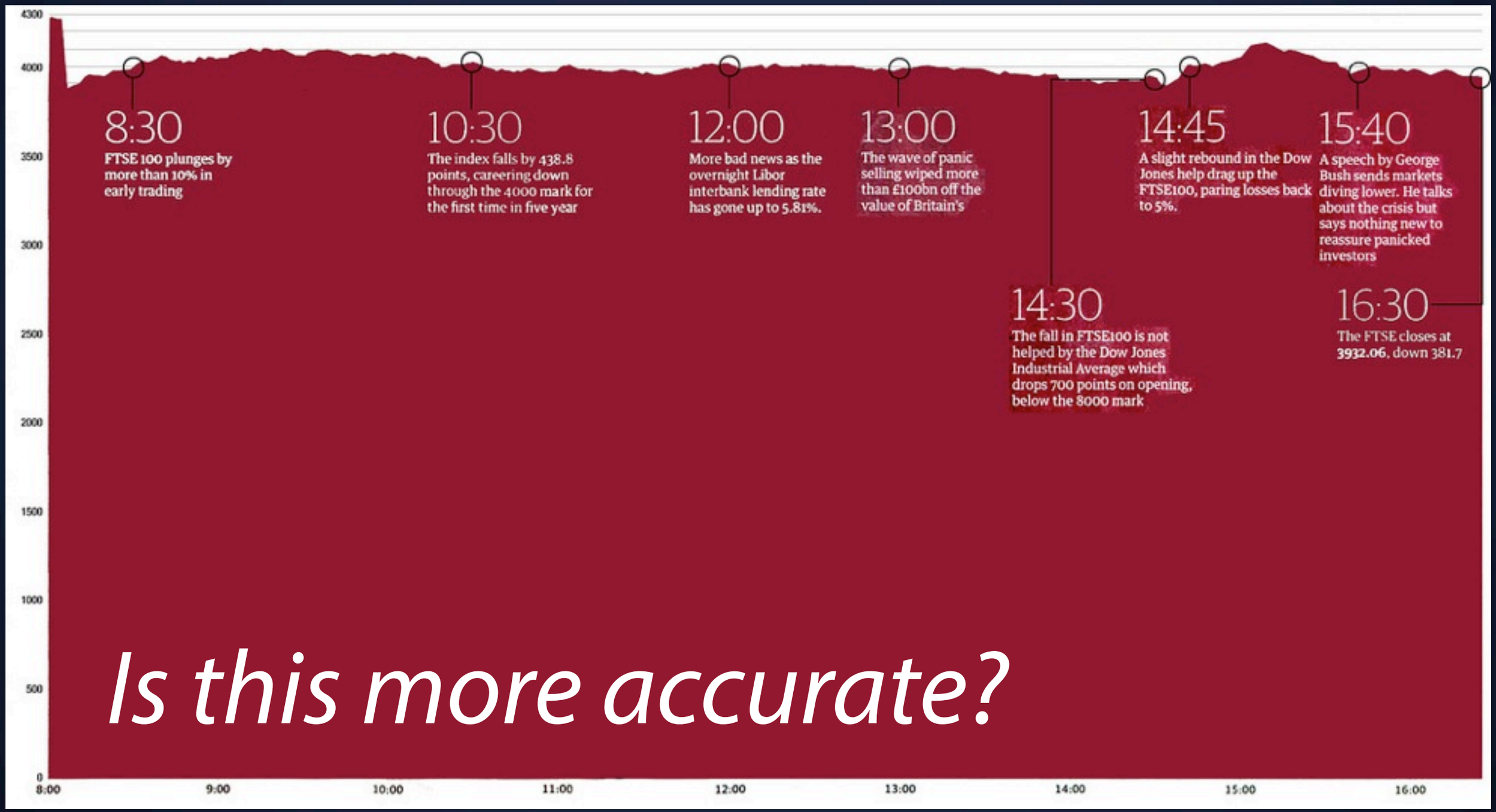
Scale / Baseline



http://www.flickr.com/photos/philgyford/2932821996/in/photostream/

Context is key

Scale / Baseline



Is this more accurate?

http://www.flickr.com/photos/philgyford/2932821996/in/photostream/

Principle 2

Show context and scale.

Indicate baseline (not necessarily at 0)

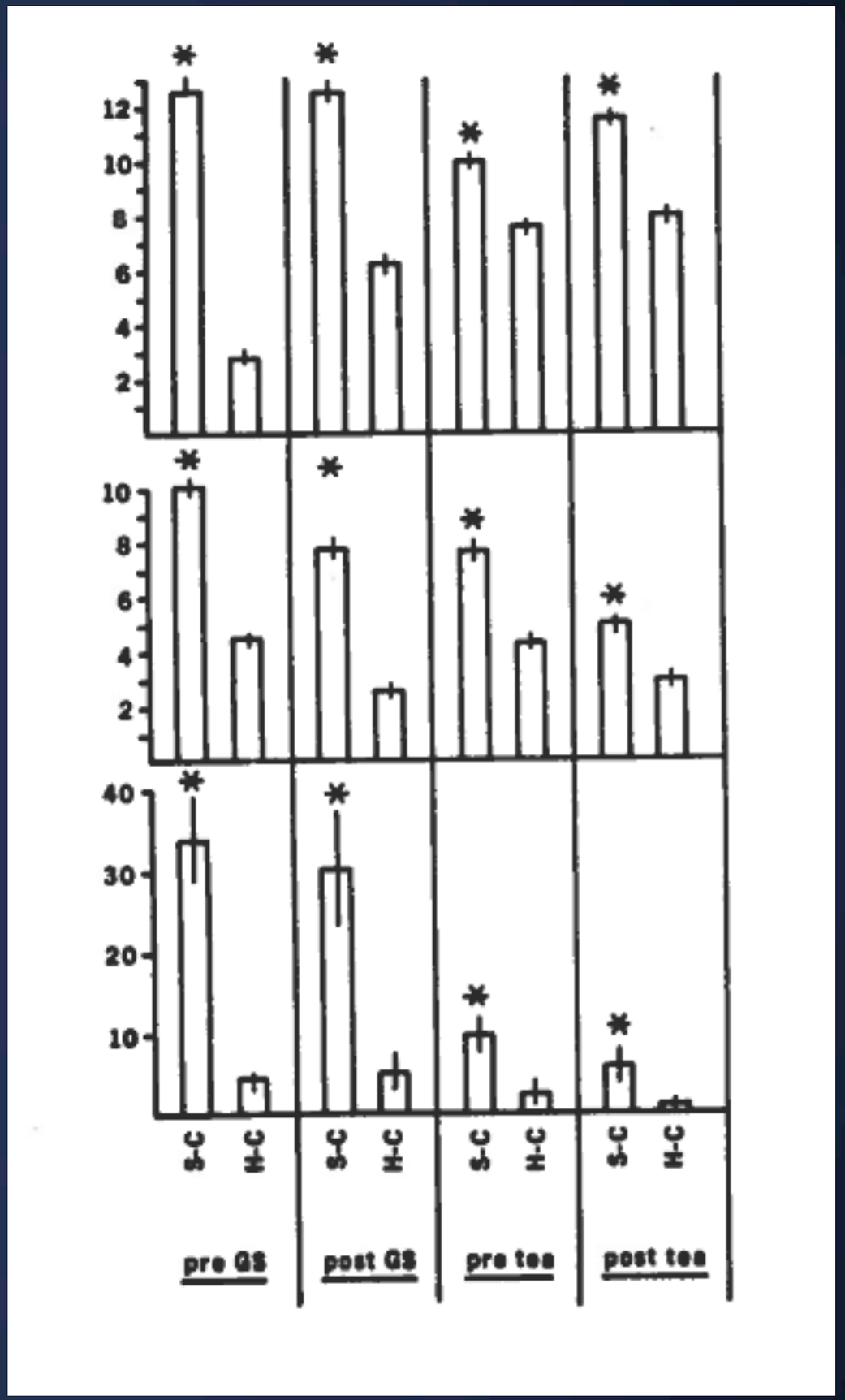
Aesthetics

...by Tufte's definition

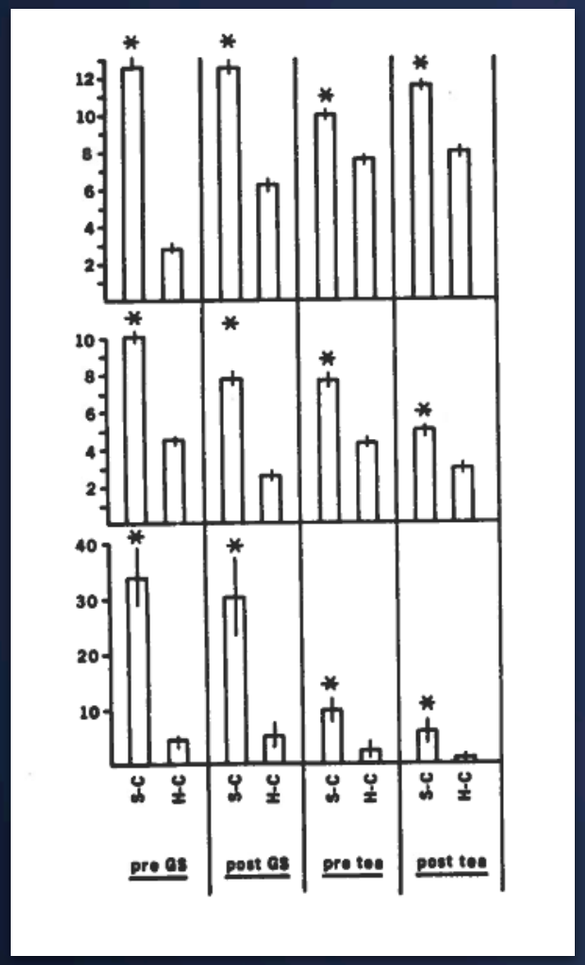
Data-Ink Ratio

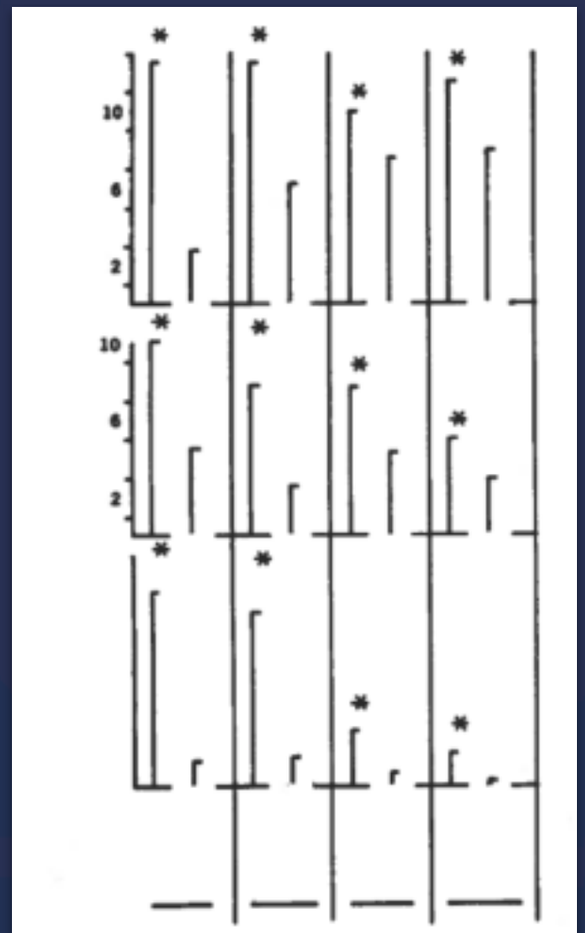
Maximizing the use of ink

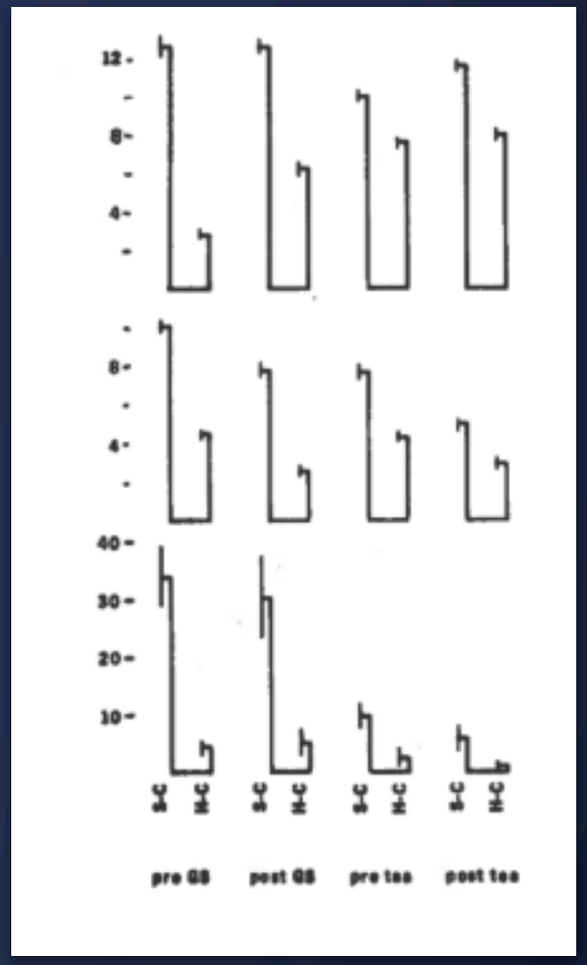
$$\textit{Data-ink ratio} = \frac{\textit{data - ink}}{\textit{total ink used}}$$



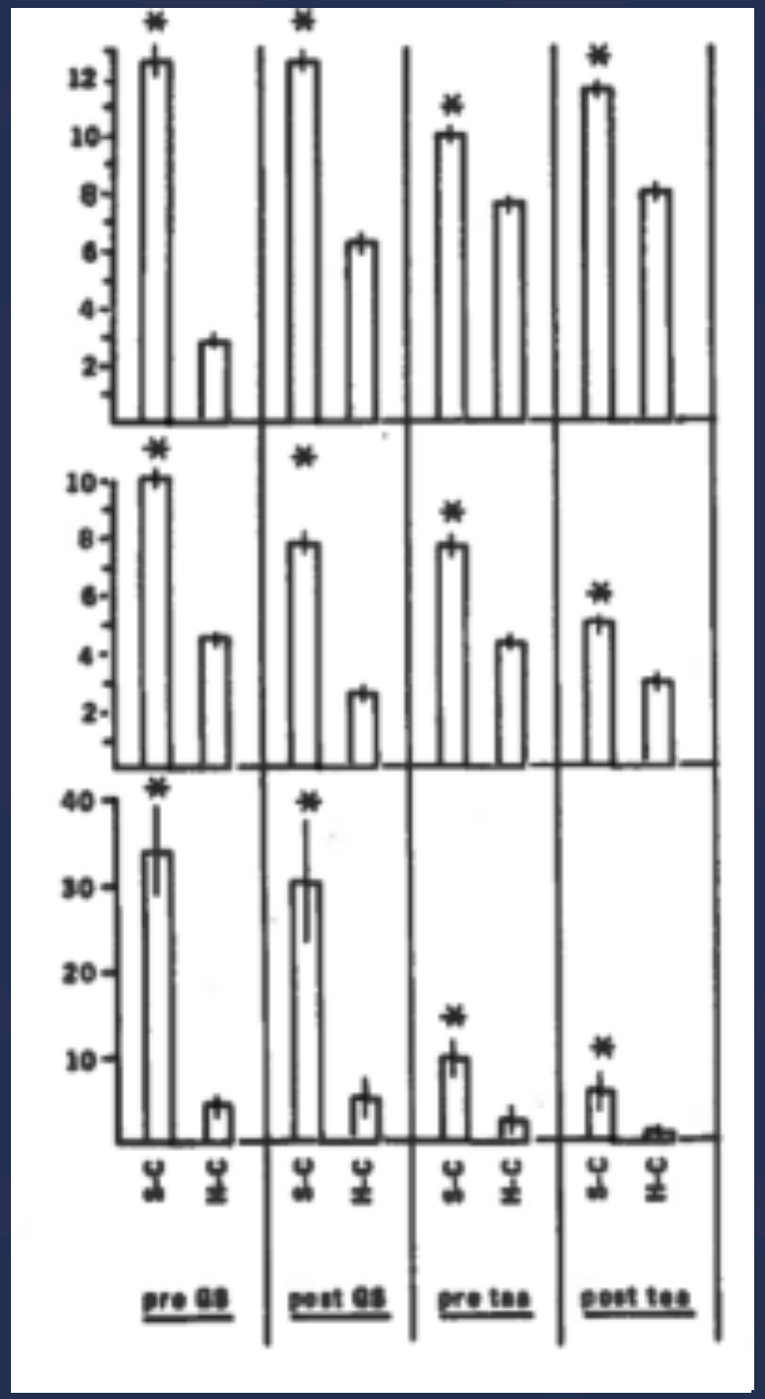
Stripping away “distraction”



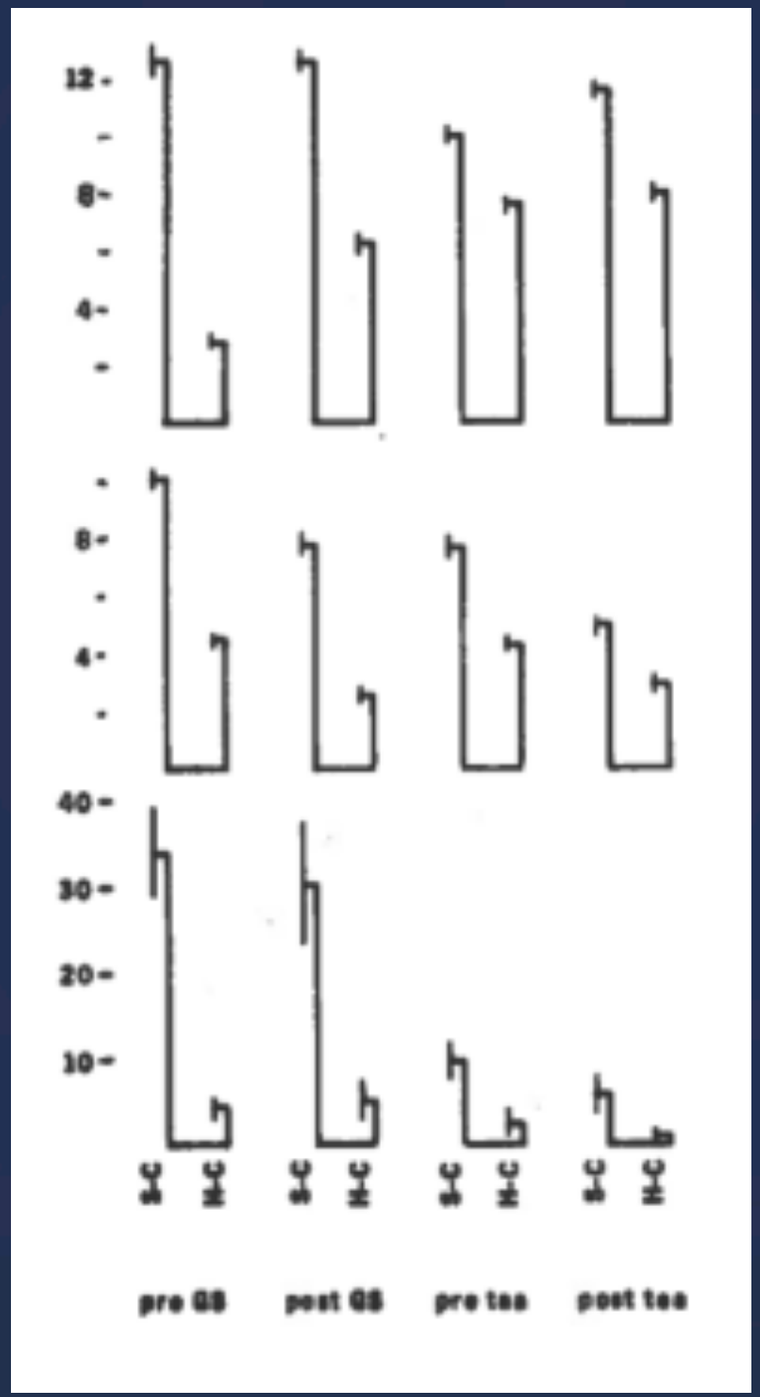




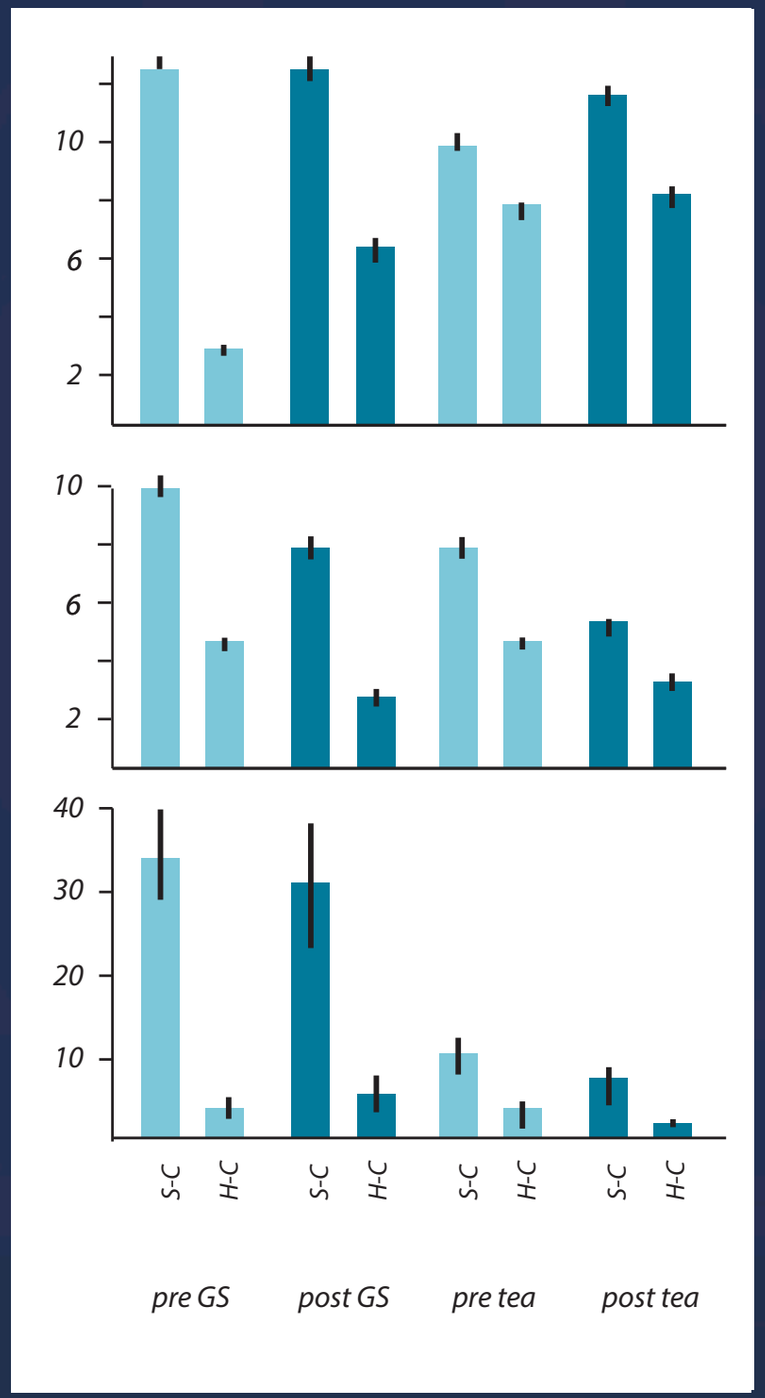
Stripping away “distraction”



Stripping away “distraction”



Stripping away “distraction”



Data-Ink

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

Data-Ink

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

Thoughts?

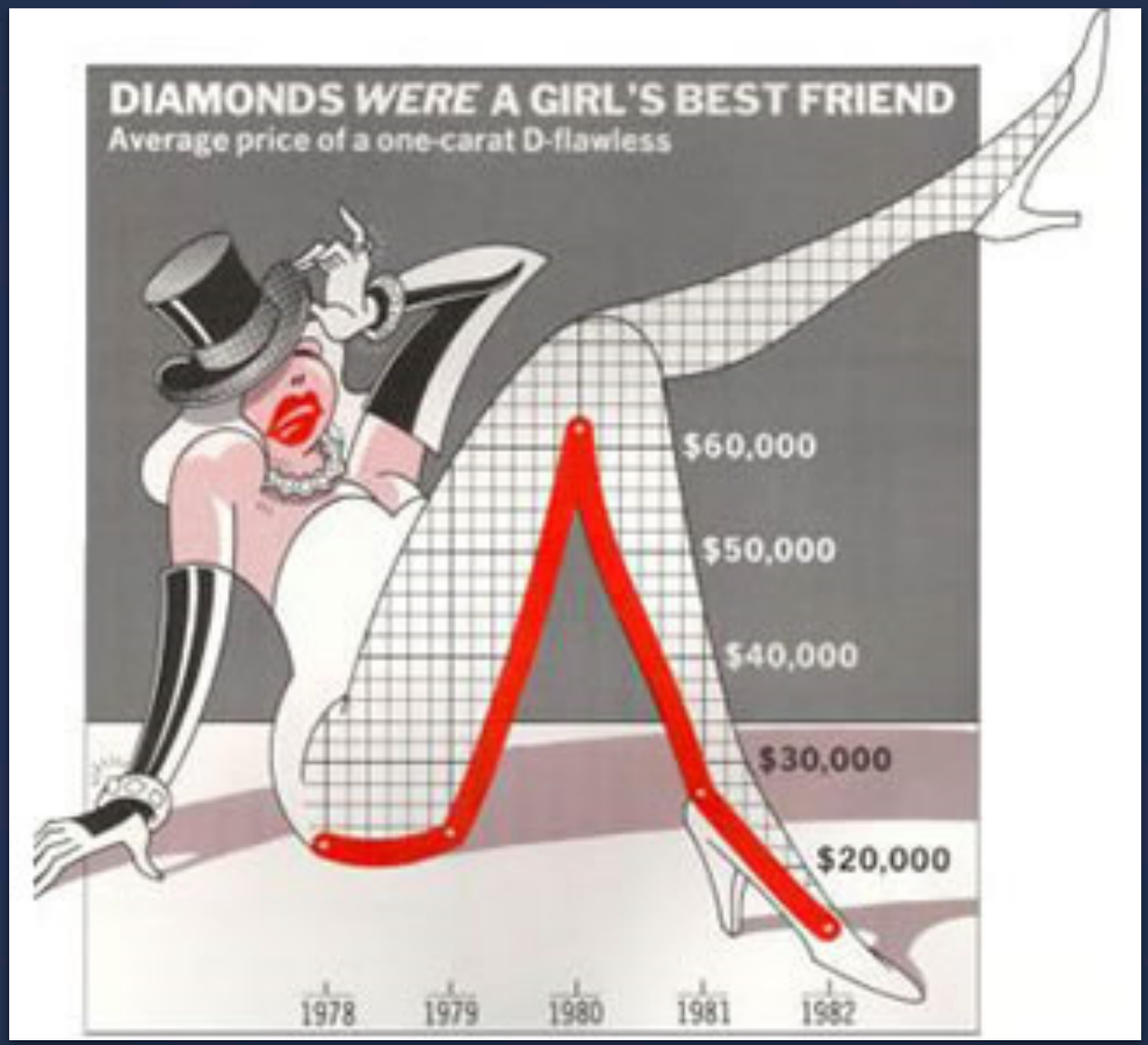
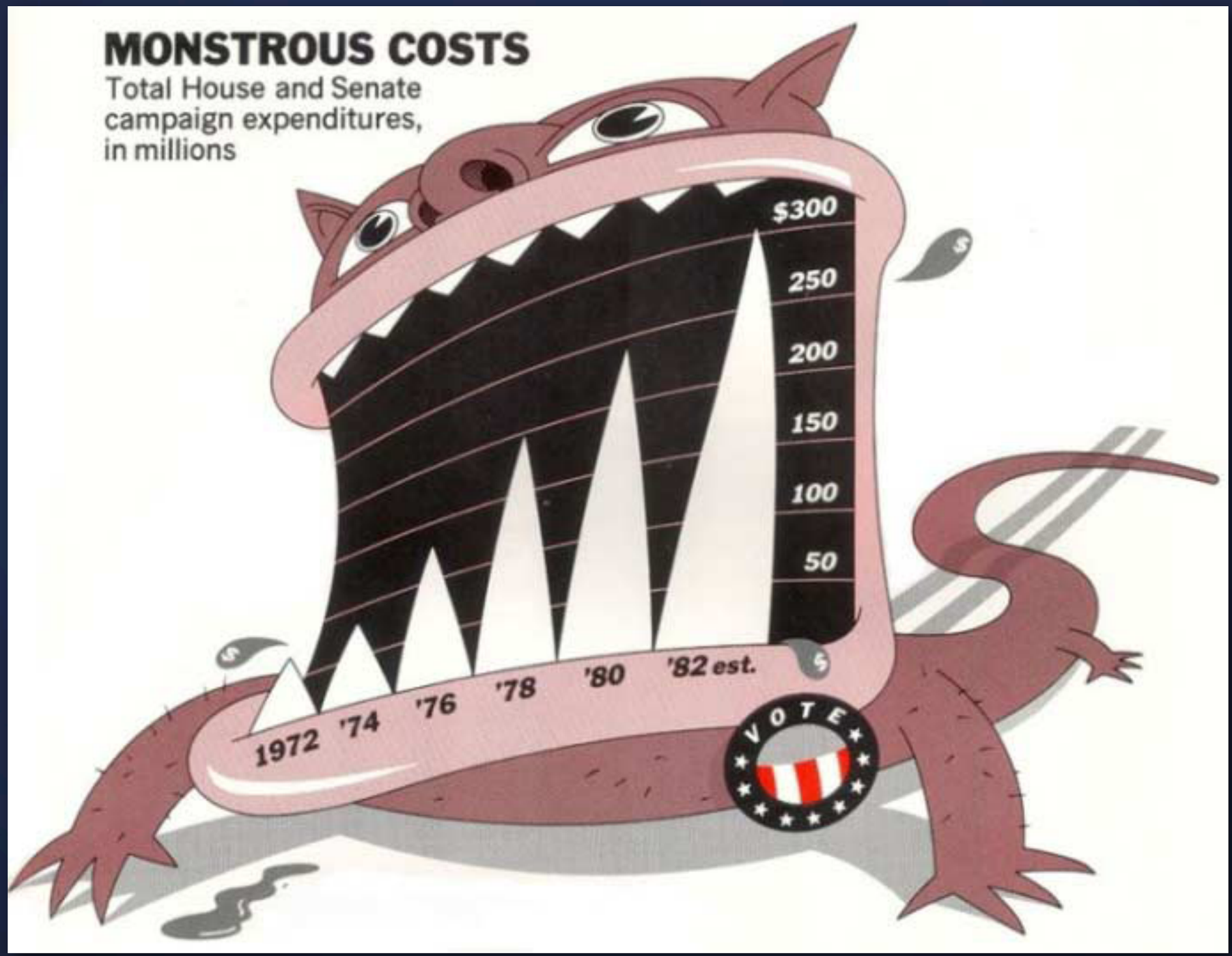

DIAMONDS *WERE* A GIRL'S BEST FRIEND
Average price of a one-carat D-flawless
$60,000
$50,000
$40,000
$30,000
$20,000
1978
1979
1980
1981
1982

Chart Junk

Chart Junk

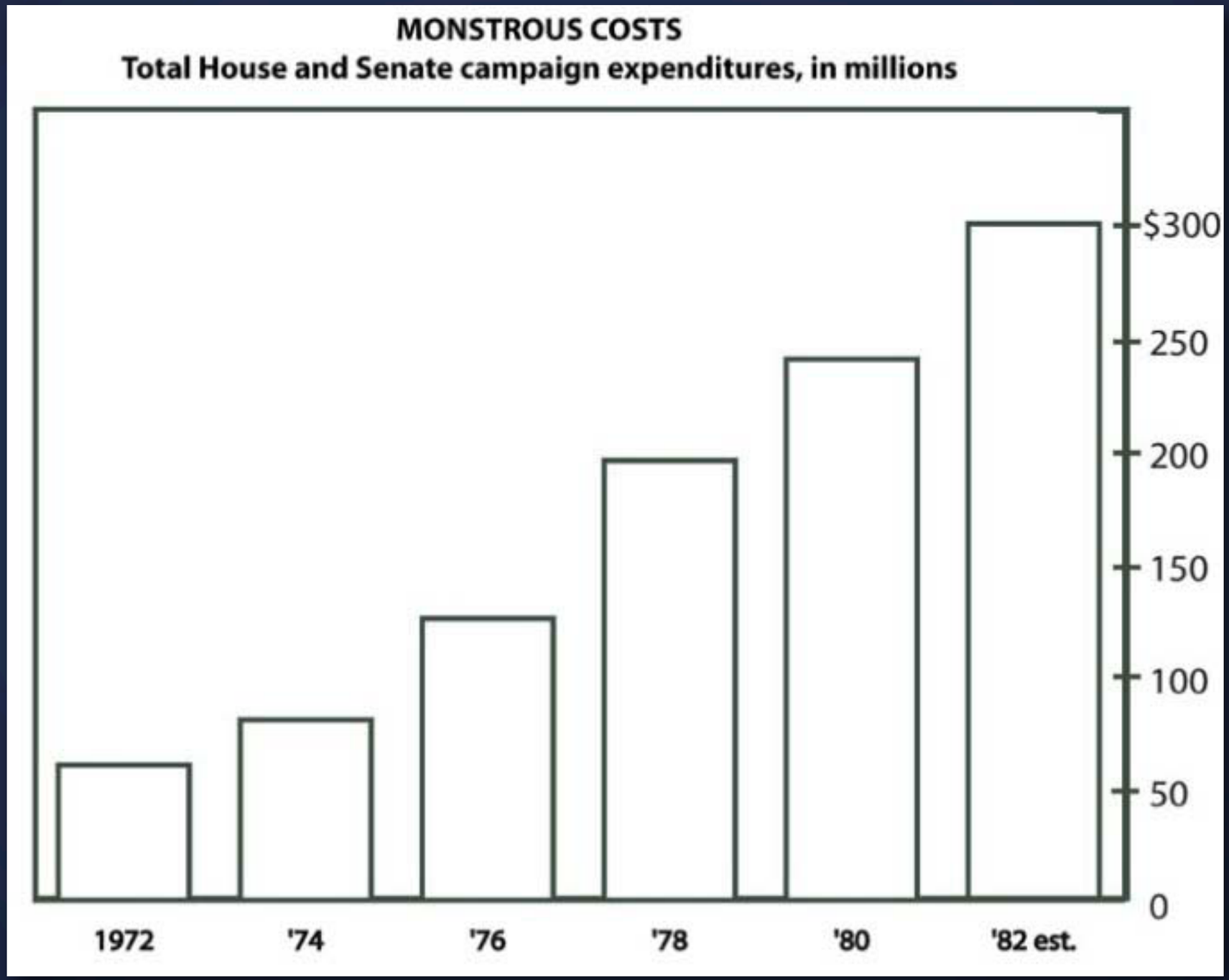
An example



examples from Bateman et al., 2010

Chart Junk

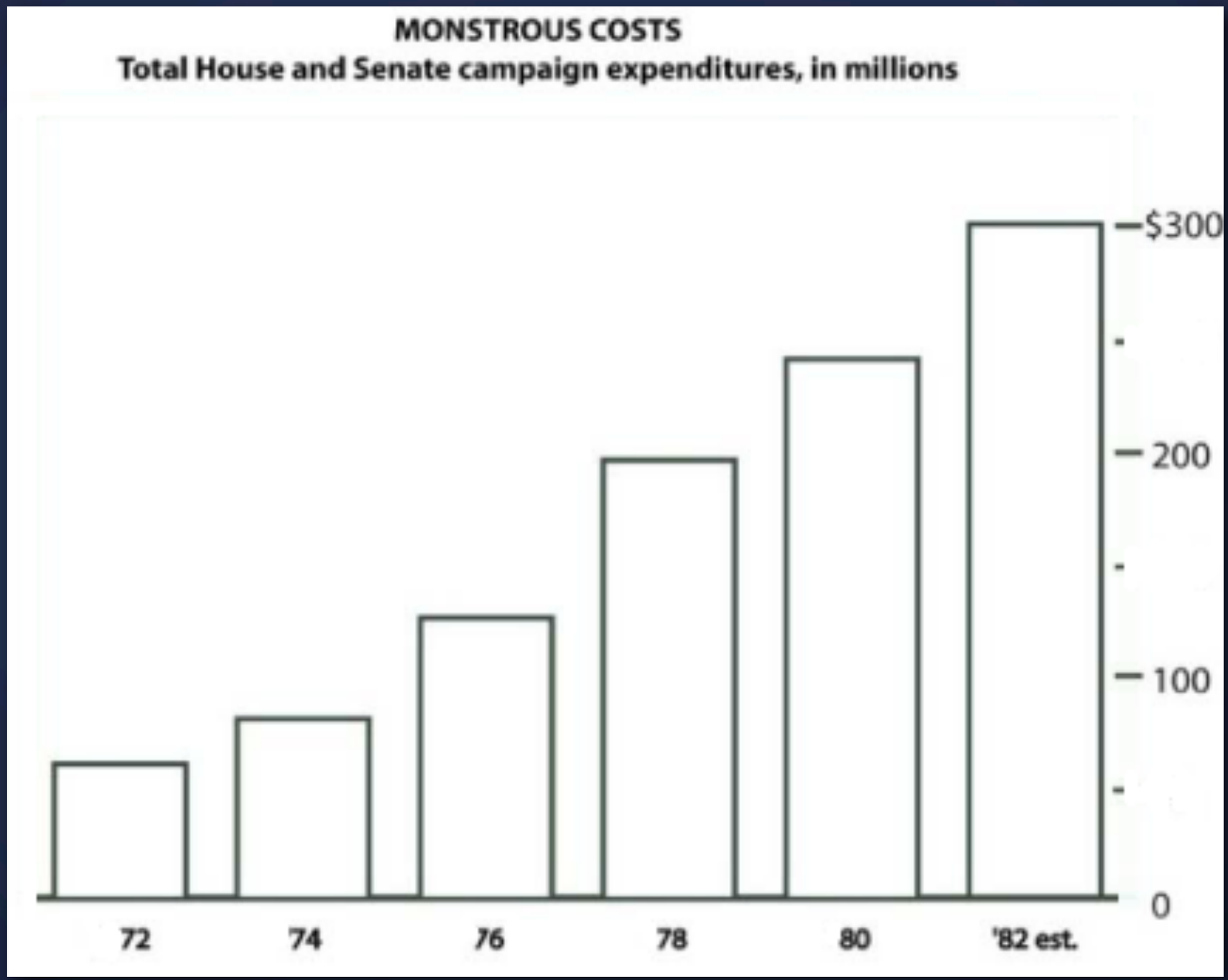
The same example



examples from Bateman et al., 2010

Chart Junk

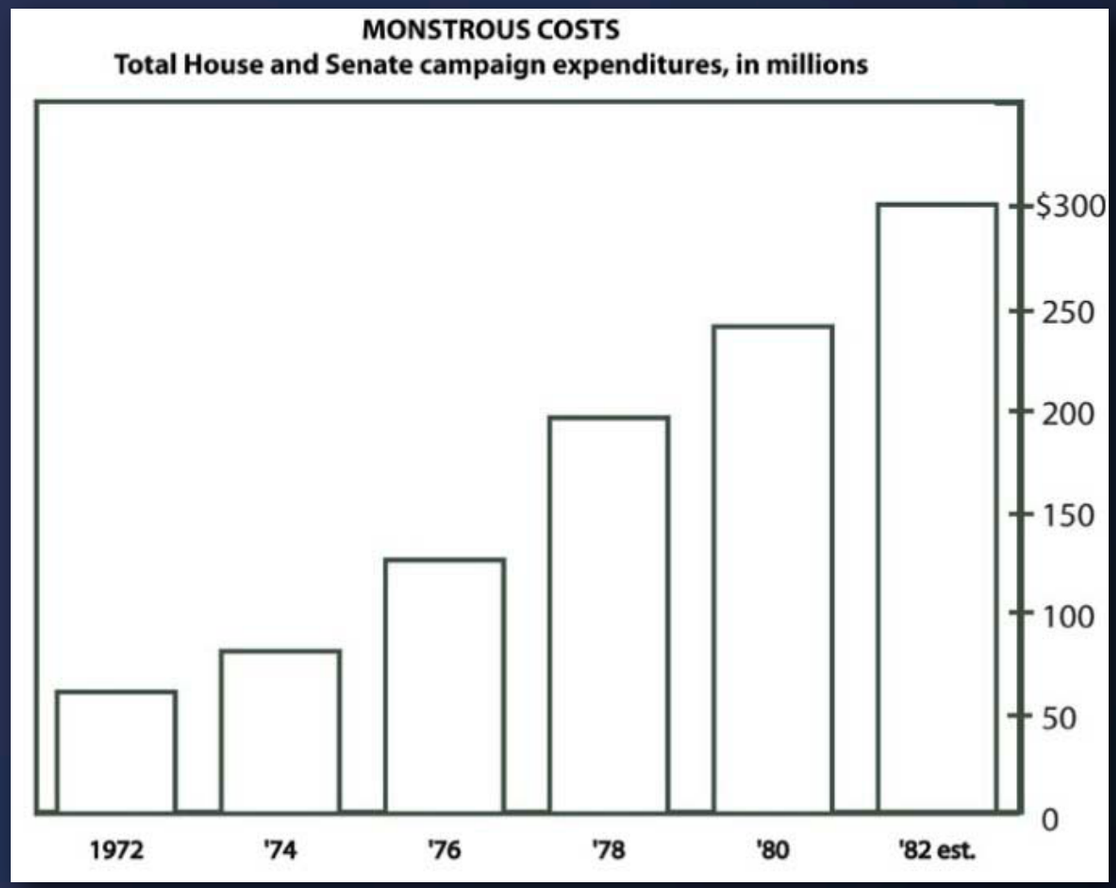
The same example



examples from Bateman et al., 2010

Chart Junk

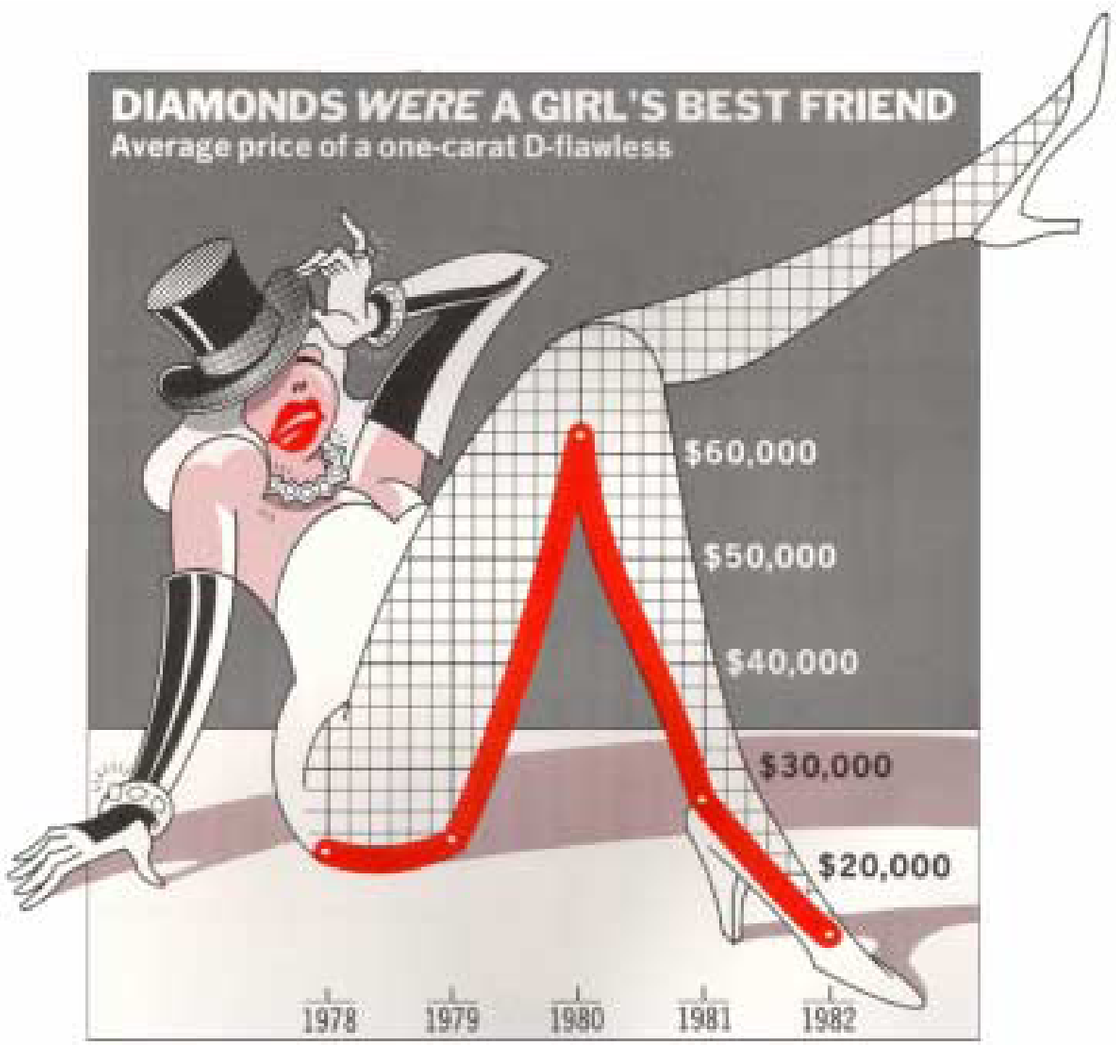
An example

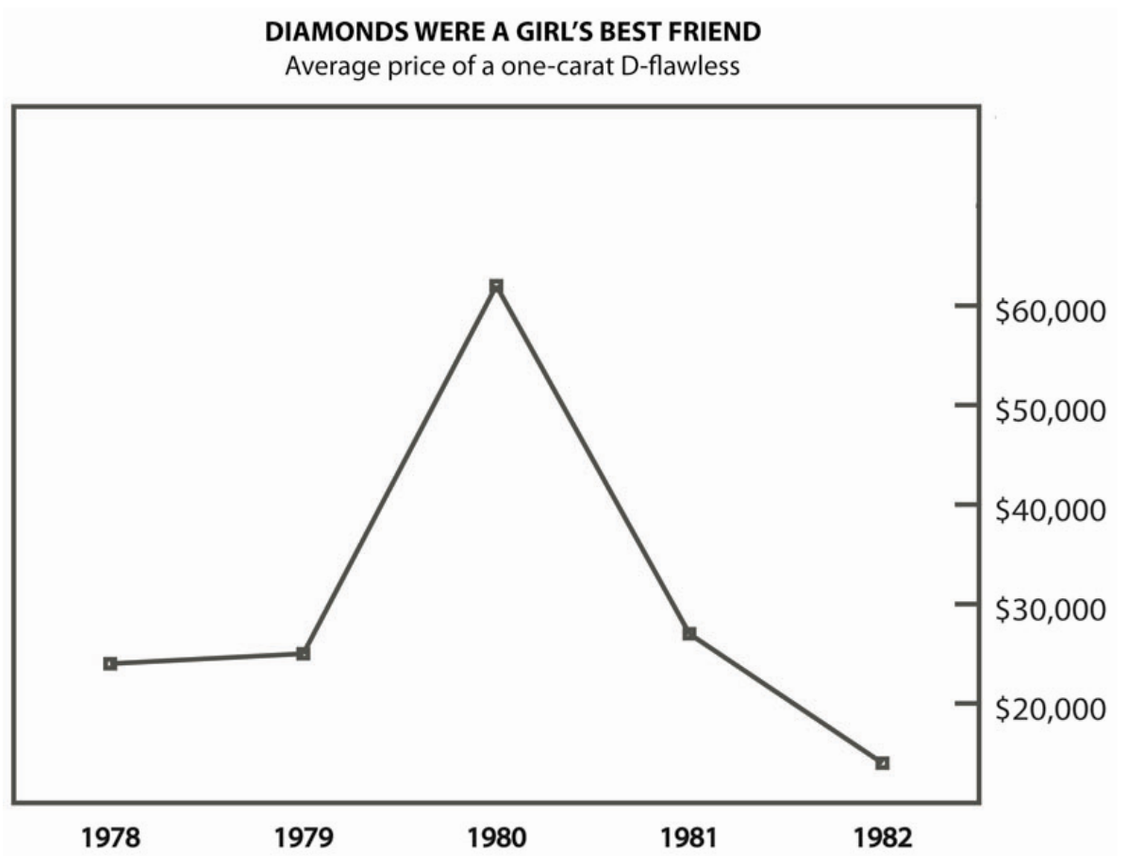


examples from Bateman et al., 2010

Chart Junk

Another example

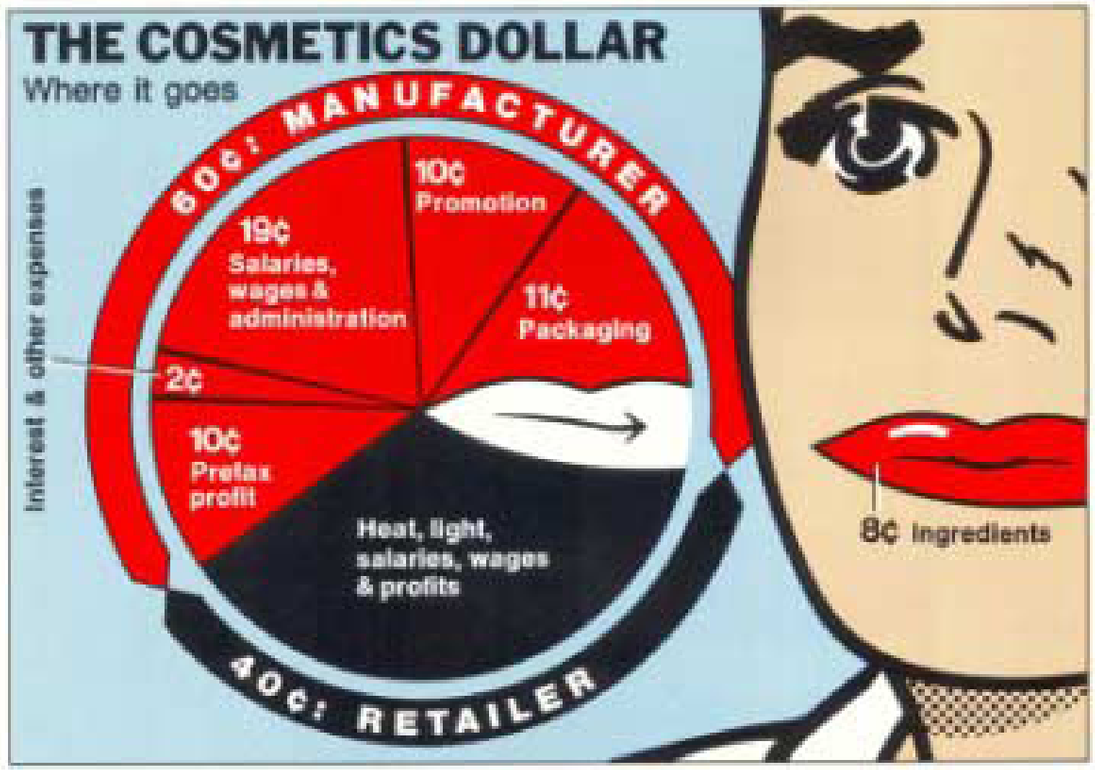


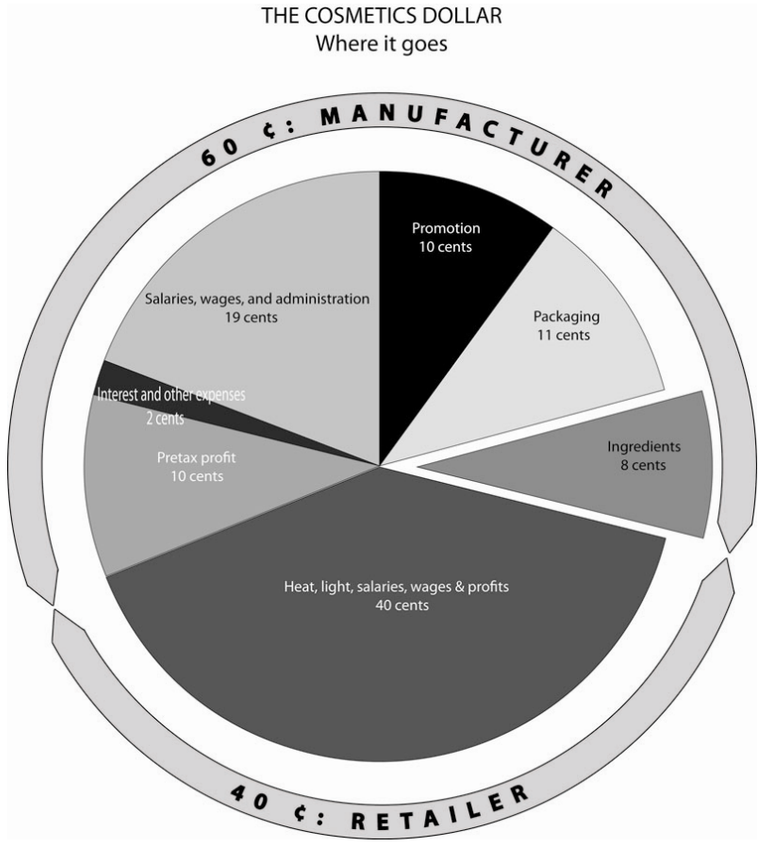


examples from Bateman et al., 2010

Chart Junk

Another example

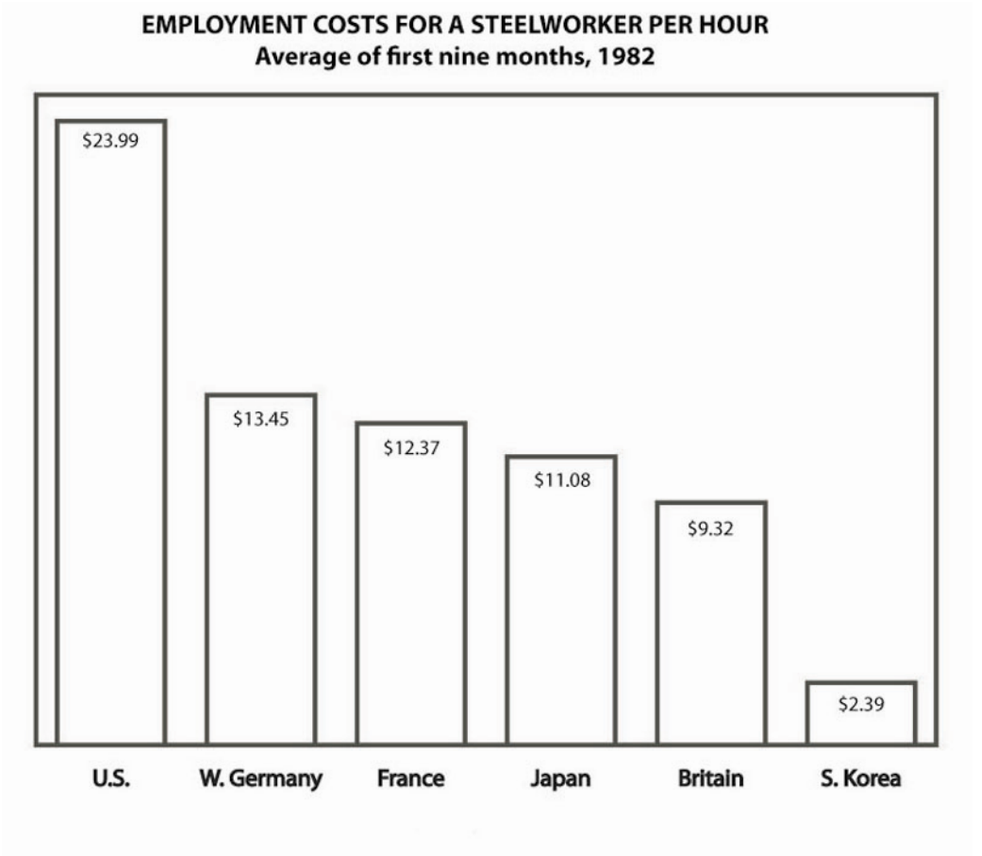




examples from Bateman et al., 2010

Chart Junk

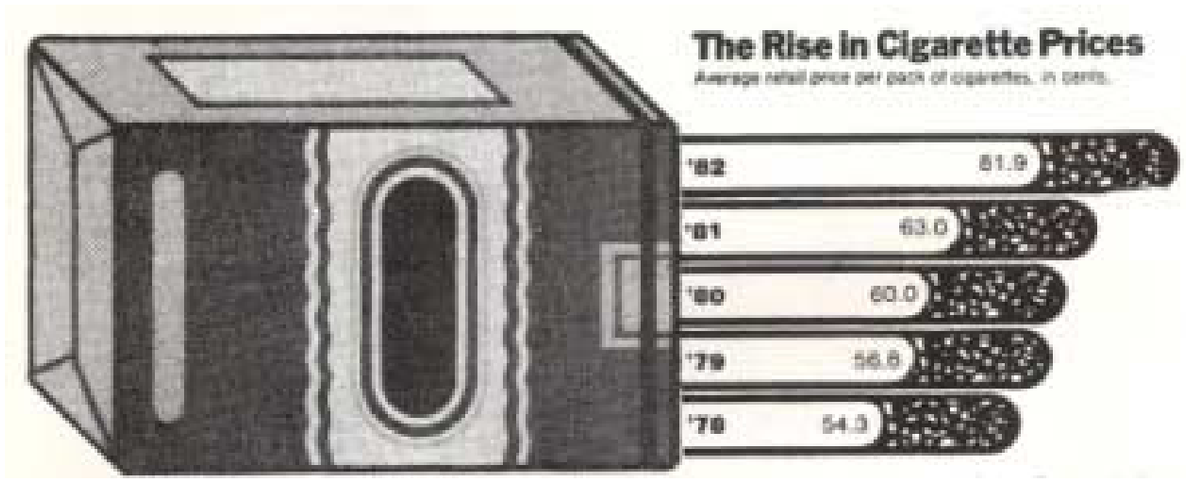
Another example

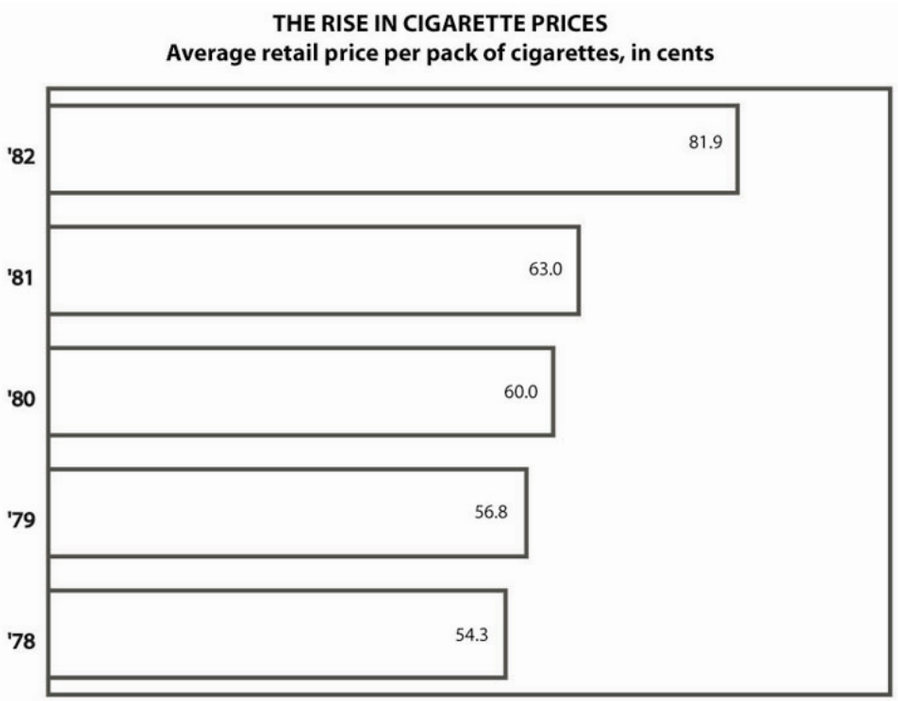


examples from Bateman et al., 2010

Chart Junk

Another example





examples from Bateman et al., 2010

Chart Junk

The Rise in Cigarette Prices

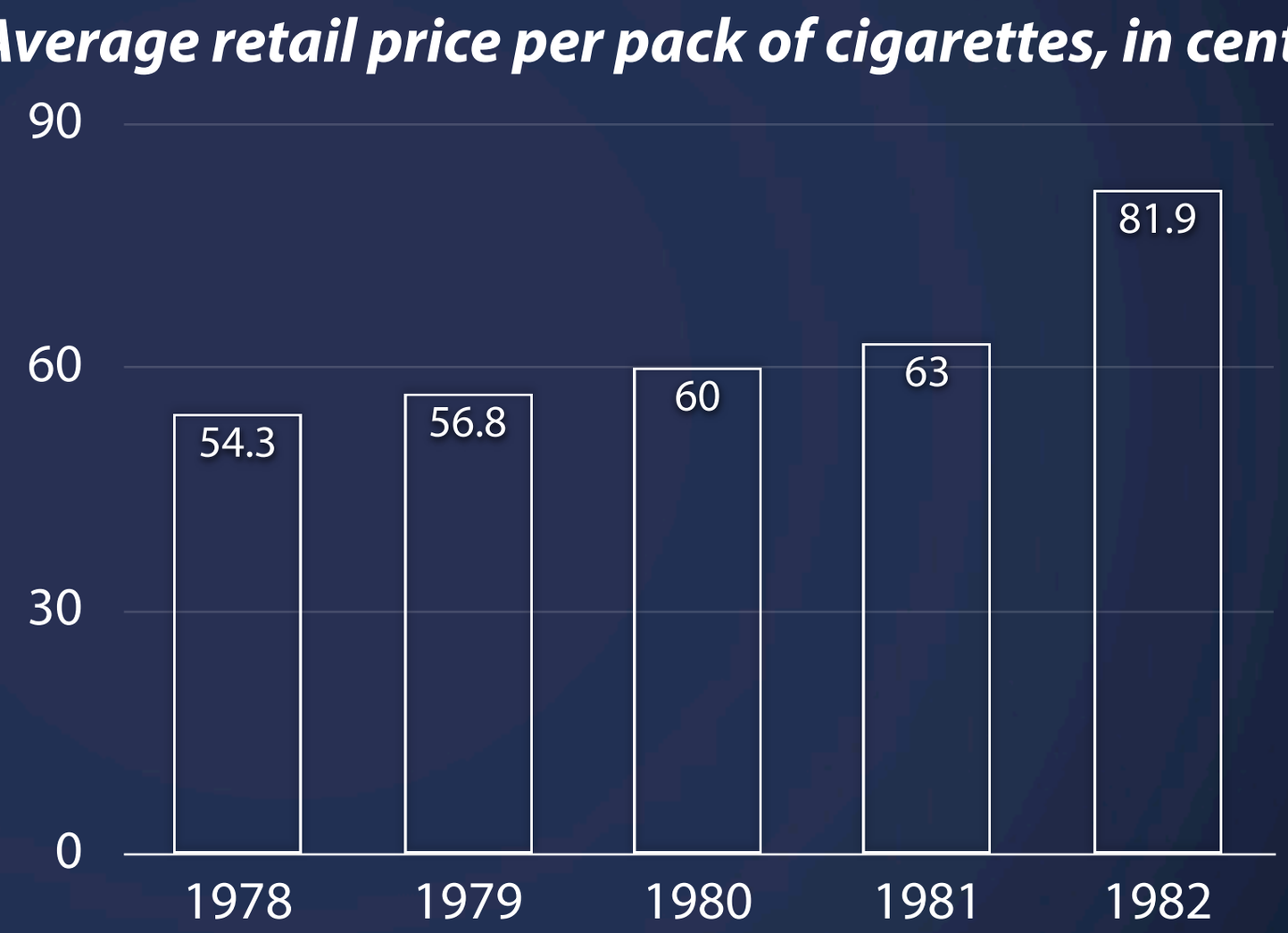


Chart Junk

The Rise in Cigarette Prices

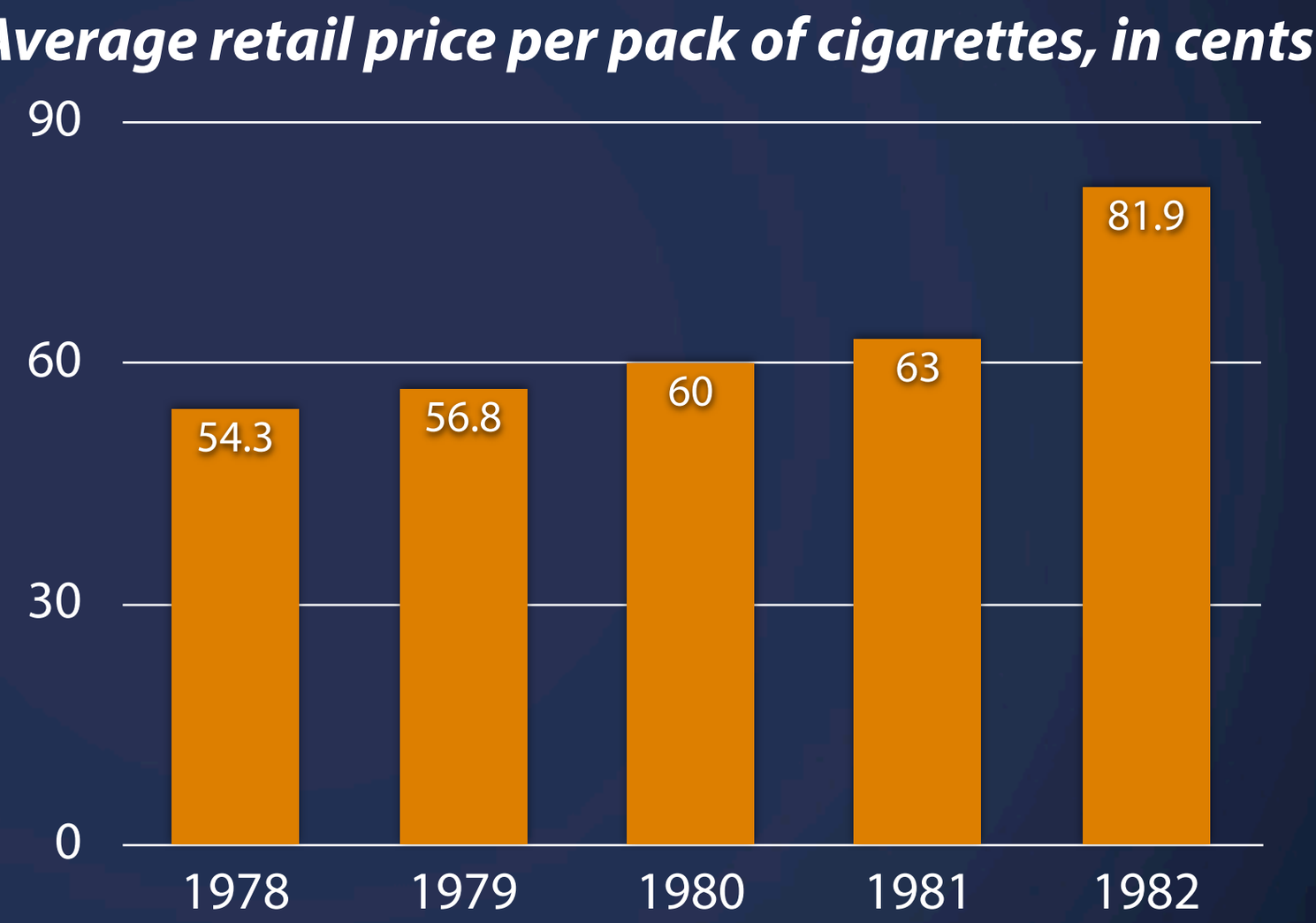


Chart Junk

The Rise in Cigarette Prices

Average retail price per pack of cigarettes, in cents

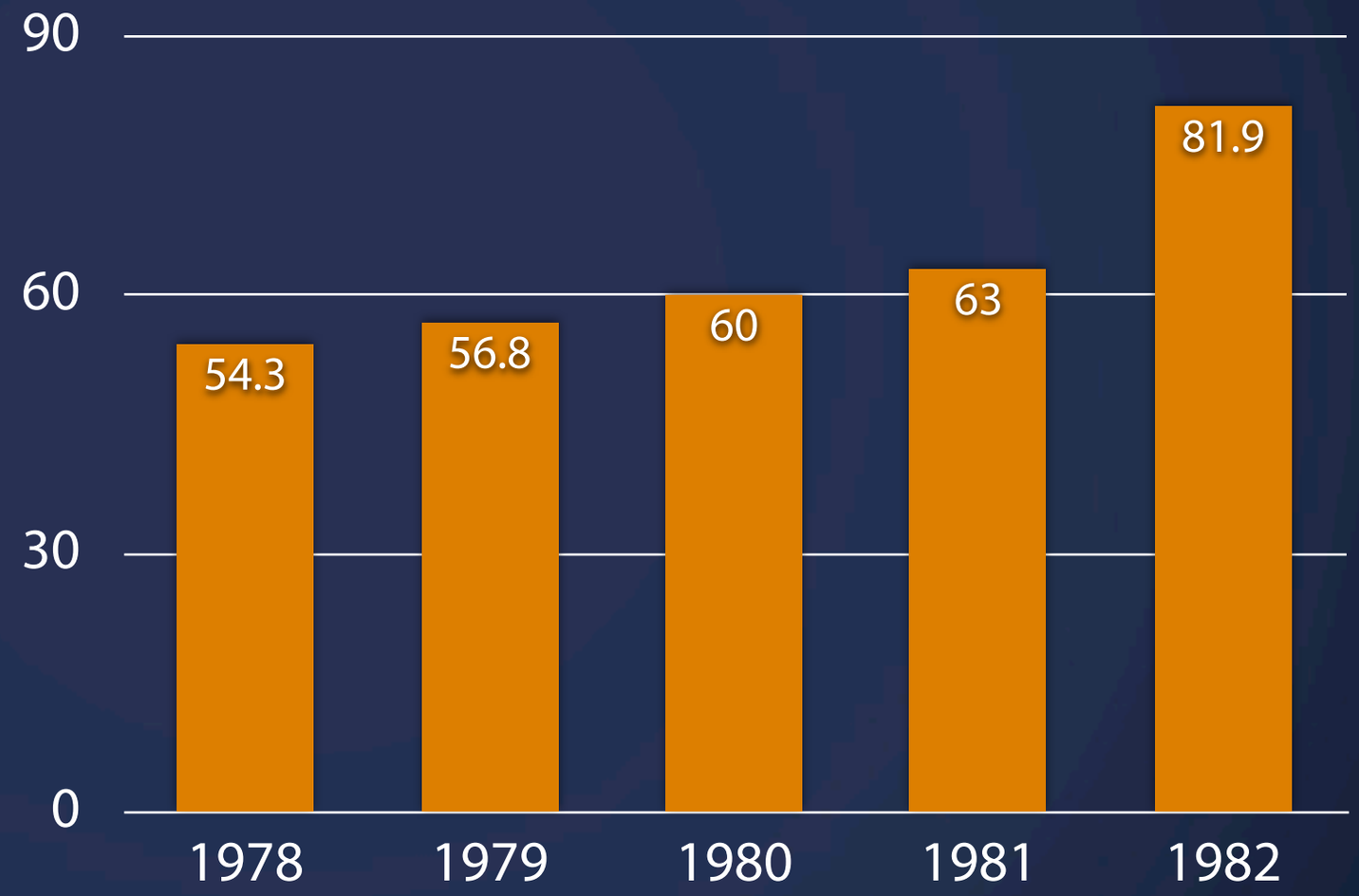


Chart Junk

The Rise in Cigarette Prices

+51%

in 4 years

Average retail price per pack of cigarettes, in cents

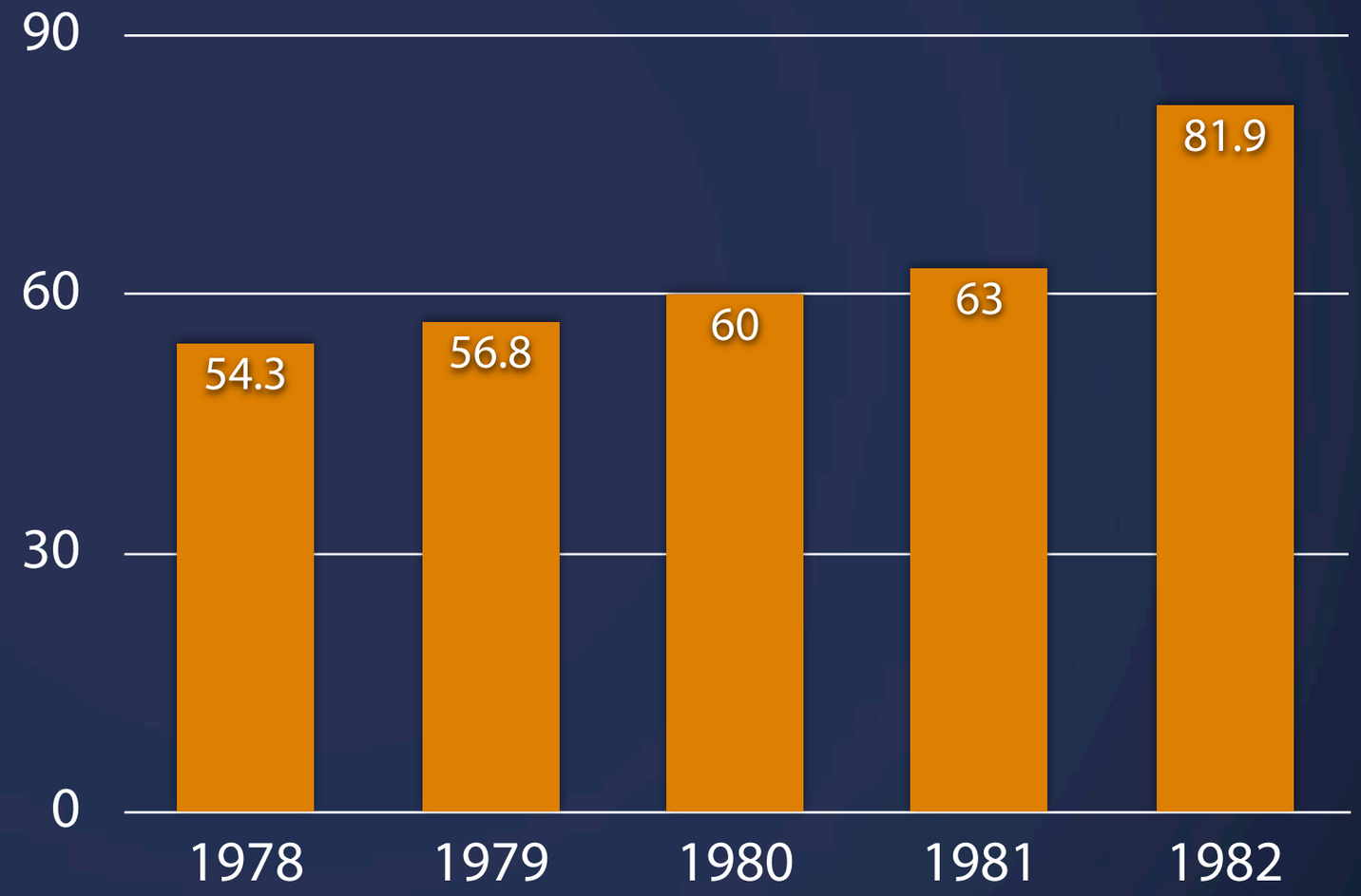


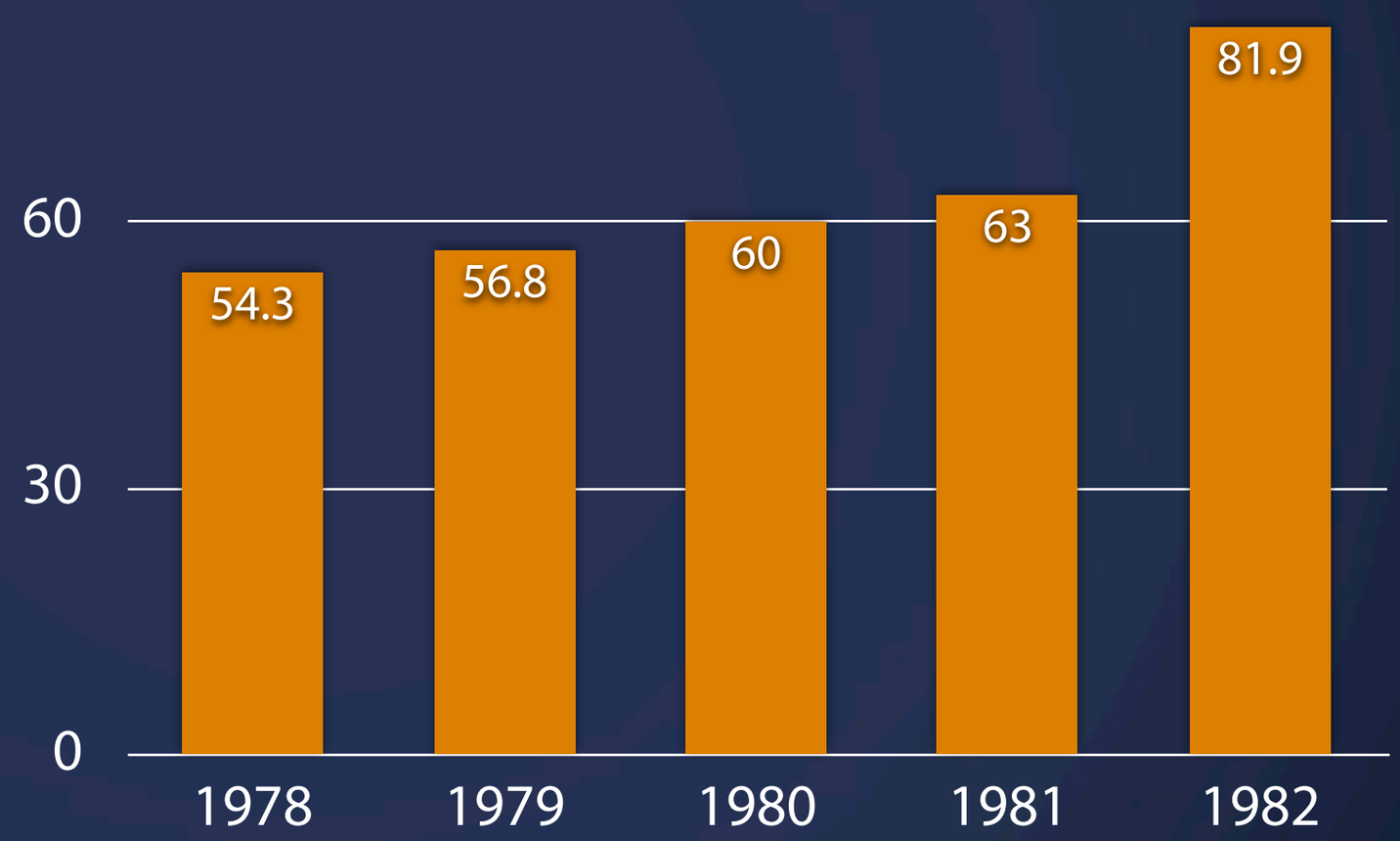
Chart Junk

The Rise in Cigarette Prices

+51%

in 4 years

Average retail price per pack of cigarettes, in cents



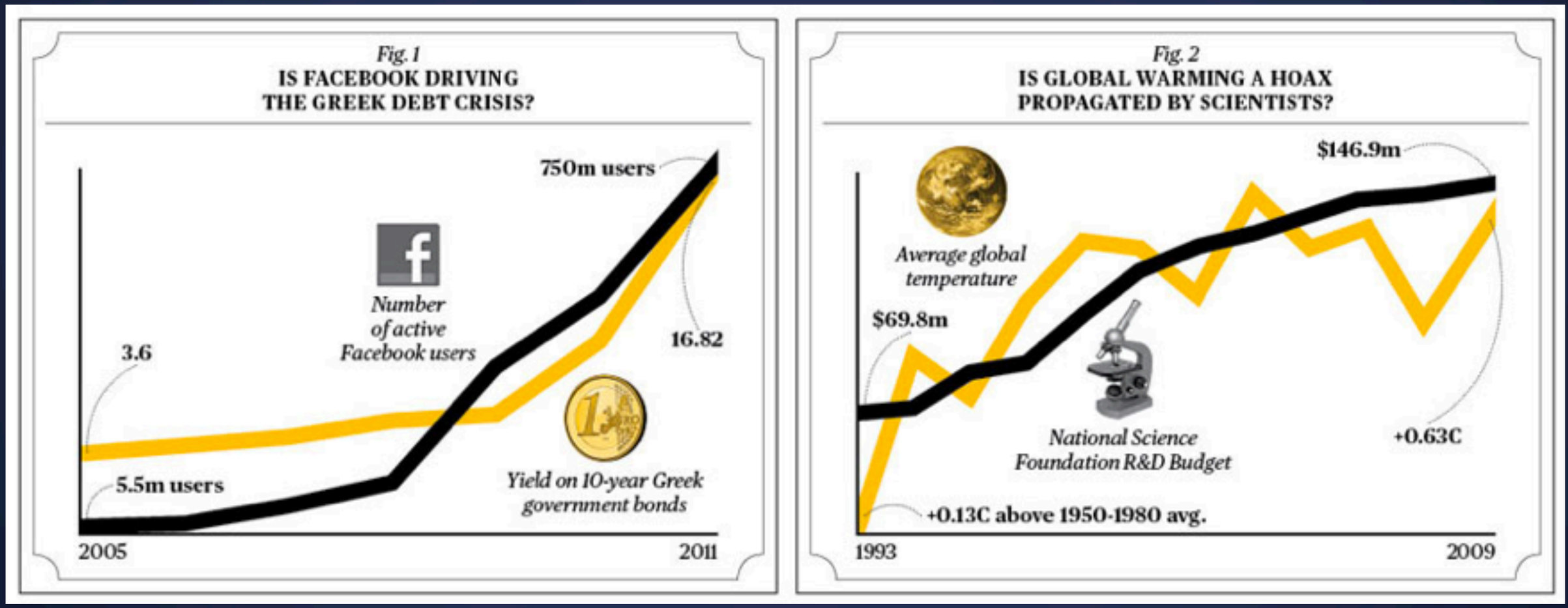
federal cigarette excise tax doubled in late 1982

Principle 3

Avoid Chart Junk, but don’t be bland

Causation / Correlation

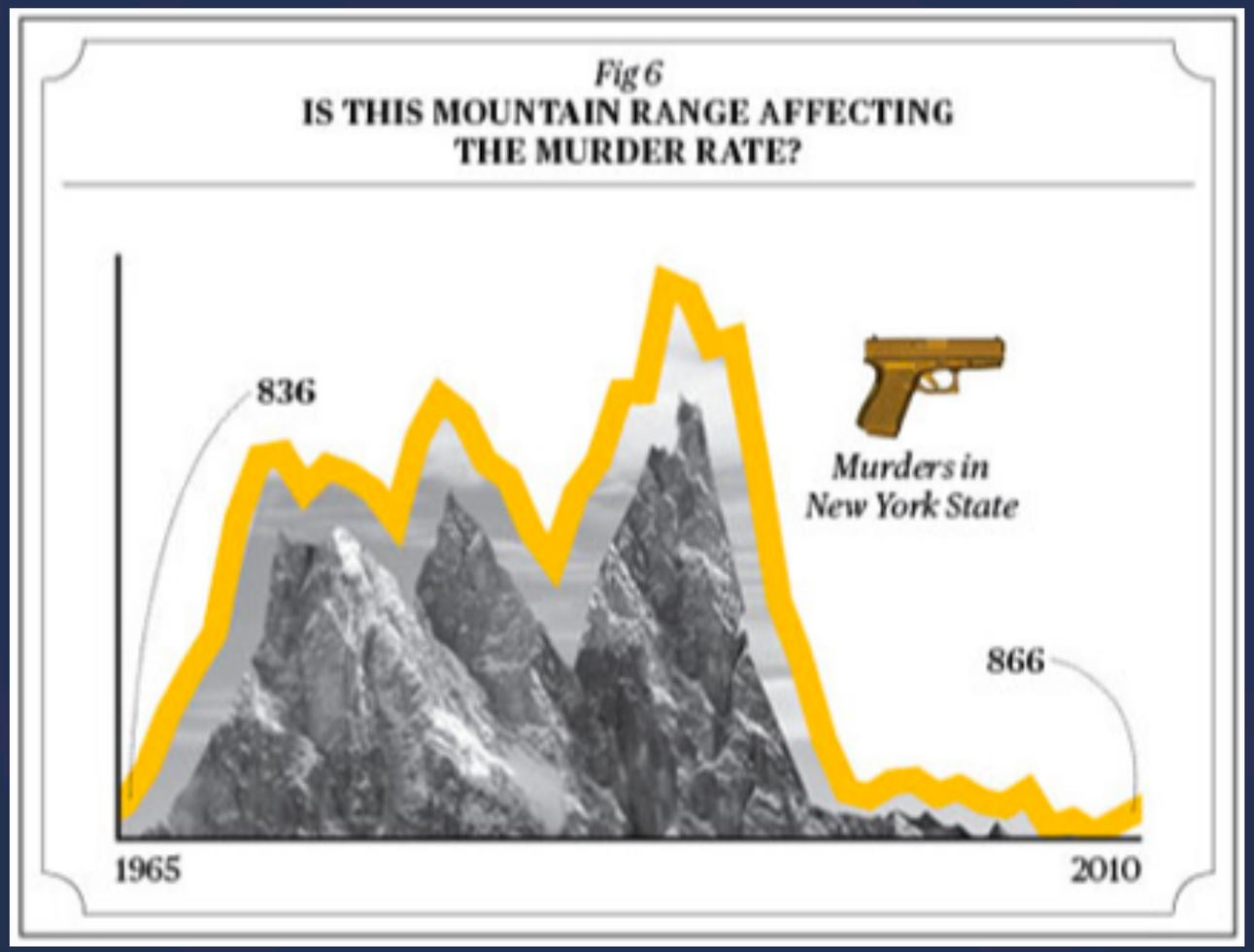
Don’t fool yourself.



http://www.businessweek.com/magazine/correlation-or-causation-12012011-gfx.html

Causation / Correlation

Don’t fool yourself



http://www.businessweek.com/magazine/correlation-or-causation-12012011-gfx.html

Principle 4

Don’t imply what isn’t

Recap

Final Project

Description and Goals

Goal

- You'll create an interactive information visualization using the tools, frameworks, and theories from this semester

- Static visualizations are accepted, if there is a rationale *not* to use interaction.

Deliverables

Writeup

- Competitive Analysis
- Sketches
- Iteration Artifacts

Blog Post

- Process and project

Final Visualization

- Hosted on a website
- Source Code

Teams

- 3-4 people
- Interdisciplinary!

Schedule

Apr 4	Project proposals due
Apr 11	**Lab:** Concept critique
Apr 25	**Lab:** Design critique
May 5	Deadline for deliverables
May 7 / May 9	Project presentations

Grading

20%	Overall quality and effort
20%	Perceptive Attributes
20%	Storytelling / Use of interactivity
20%	Effective representation / encoding of data
20%	Process / Iterations

Project Proposals

- 1-2 pages
- Who is in your team
- The topic of your visualization
- Everything you already know: data sets, audience, etc
- 2 examples from competitive analysis
- Due **Thursday, April 4, 3PM**