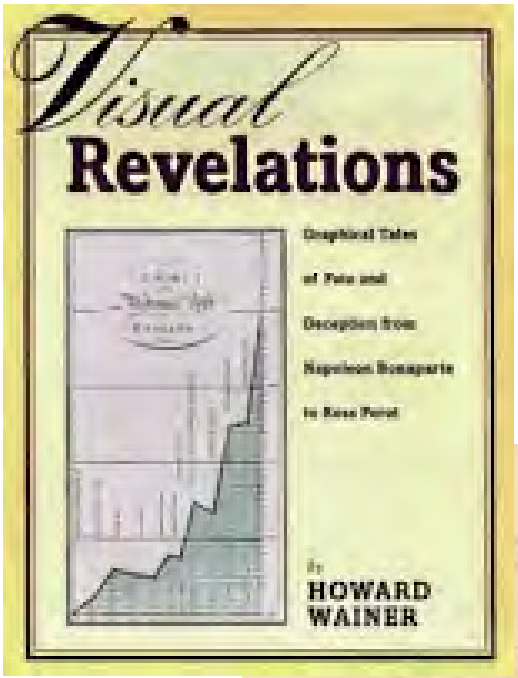
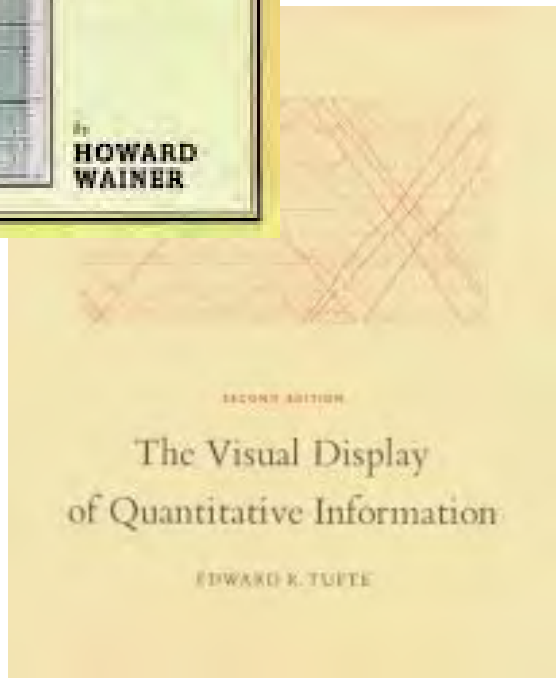
Correcting graphical lies



Howard Wainer



Edward Tufte

and others

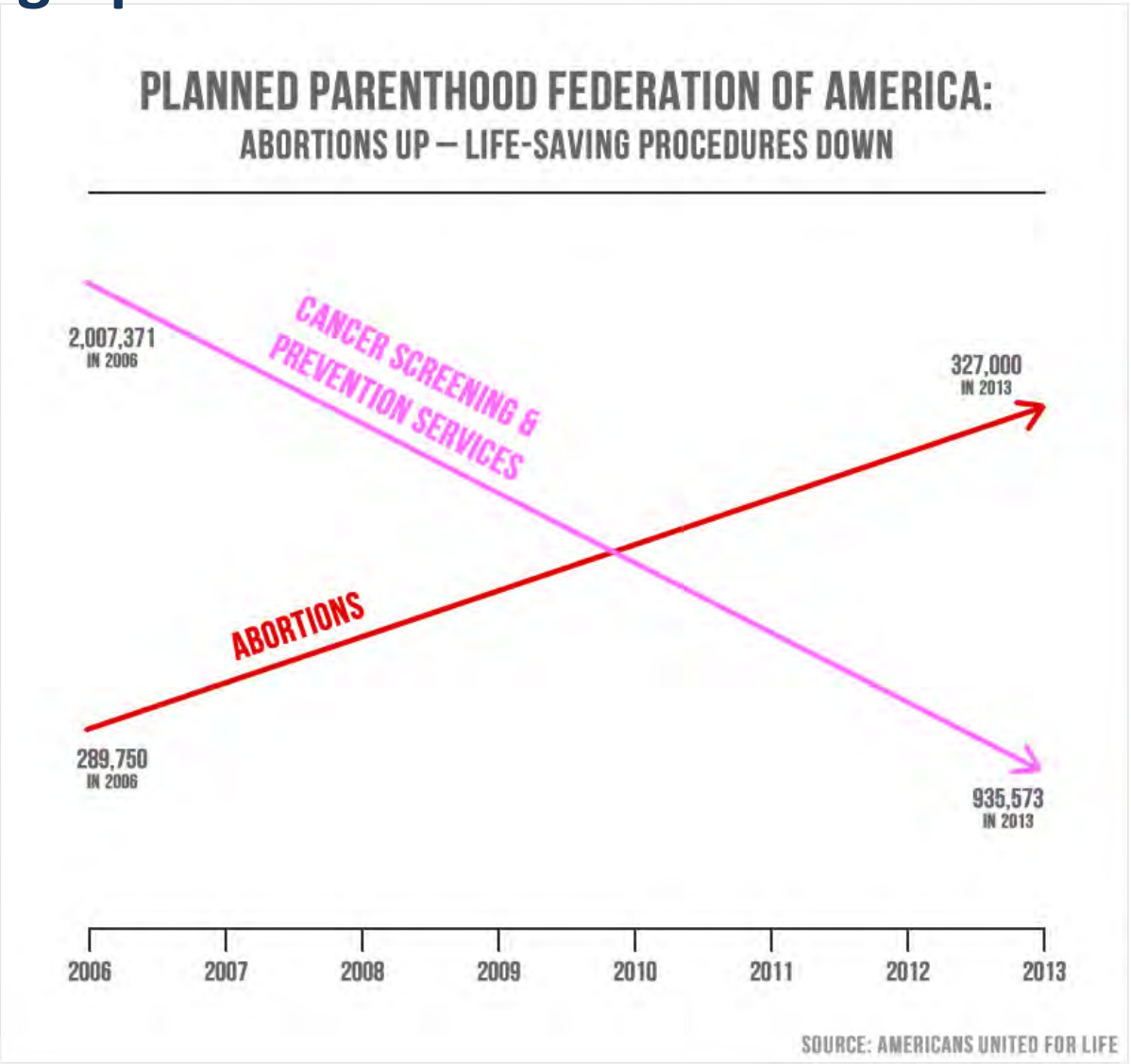
a deliberate deception

visual lie
missing data
uses emotion-laden color
lack of context
not adjusted for population

http://www.graphdoctor.com/archives/1248

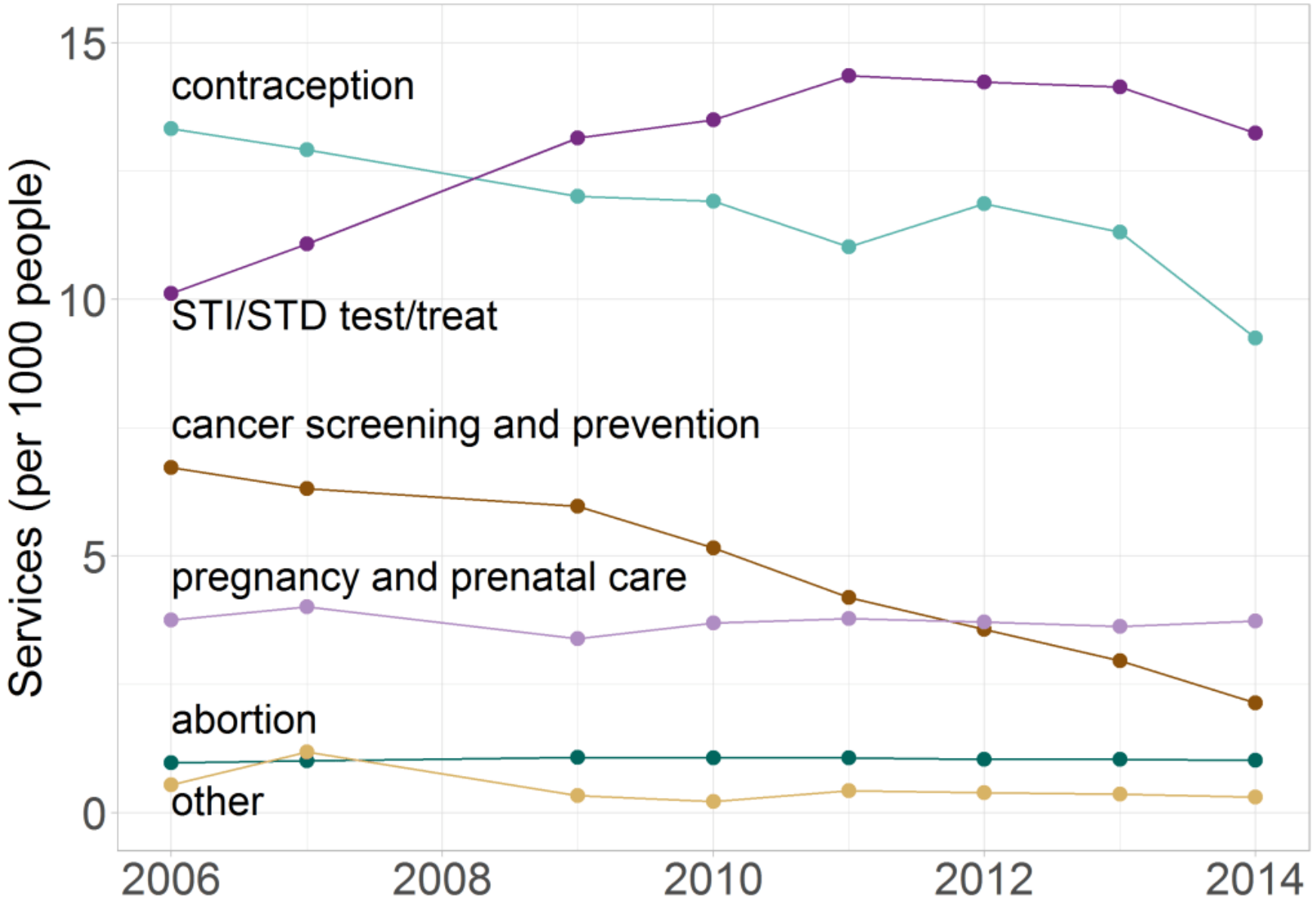
Sep. 29, 2015, Rep. Jason Chaffetz supported his assertion with this graph.

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



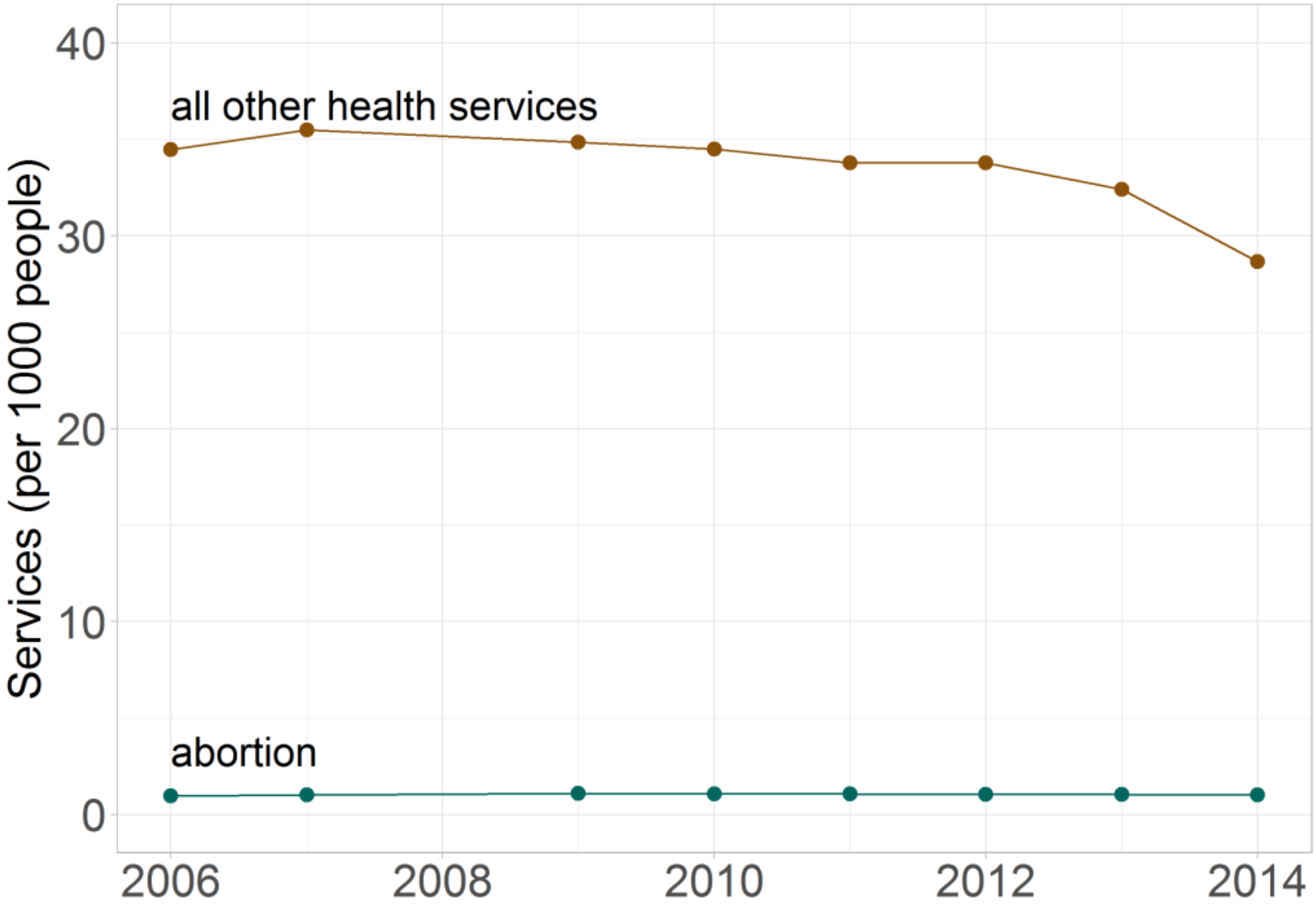
Linda Qui (2015-10-01) Chart shown at Planned Parenthood hearing is misleading and 'ethically wrong' www.politifact.com

All Planned Parenthood services, per capita



data source: Annual Reports, Planned Parenthood, 2006–2014

Abortion services compared to all other services

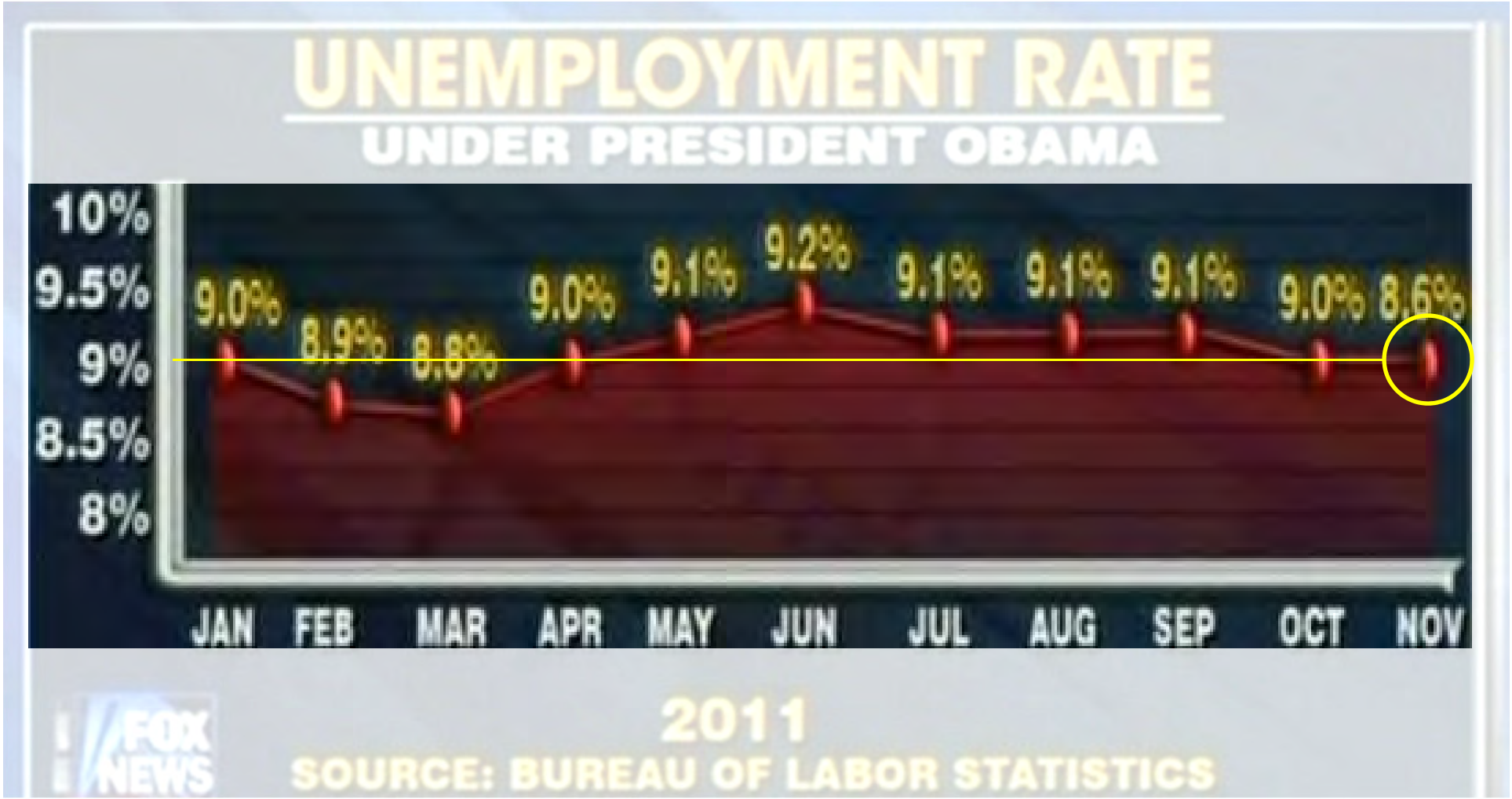


data source: Annual Reports, Planned Parenthood, 2006–2014

a deliberate deception

visual lie
lack of context

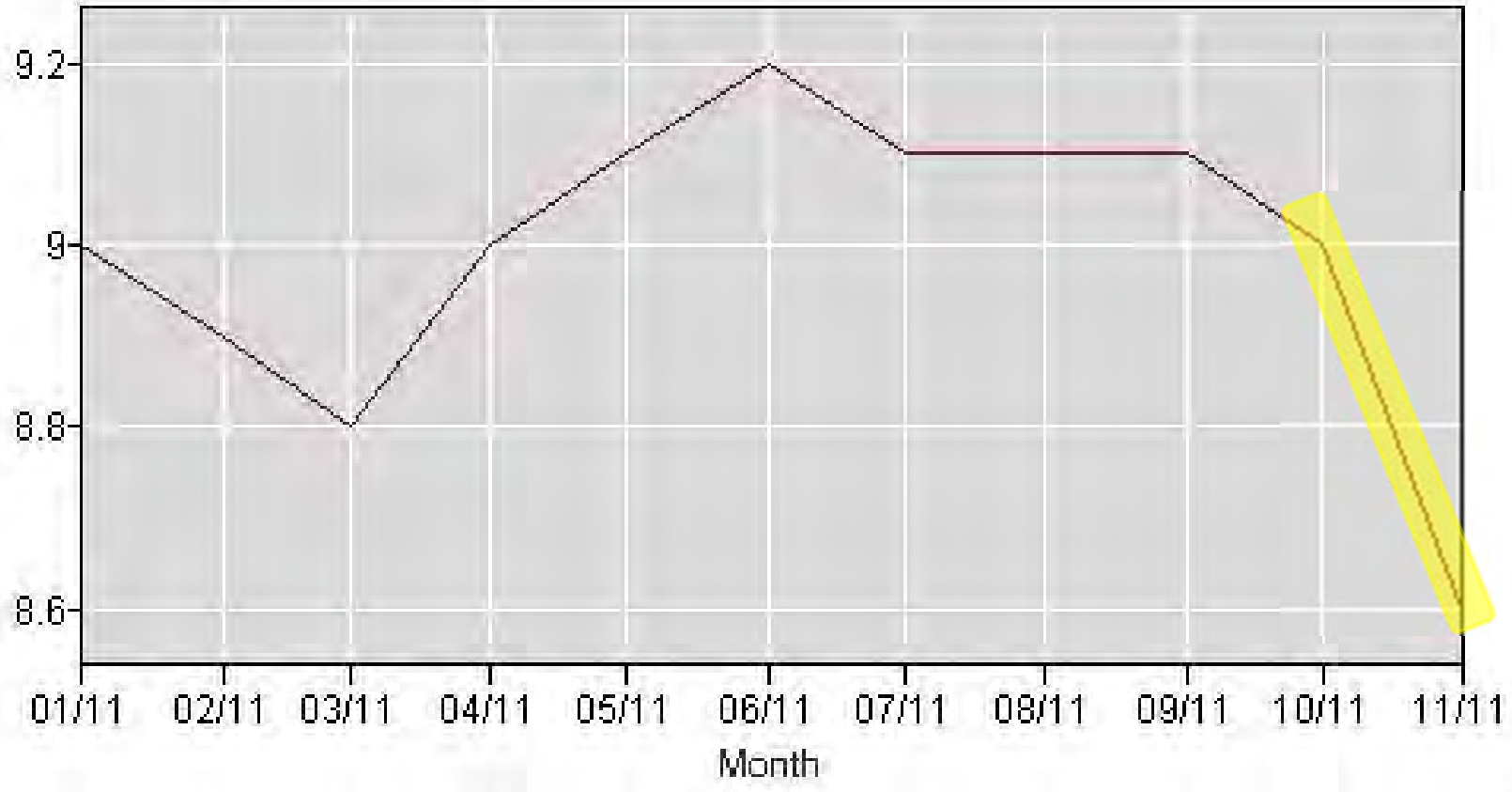
The November data label (8.6%) is correct. What is amiss?



http://mediamatters.org/research/2012/10/01/a-history-of-dishonest-fox-charts/190225

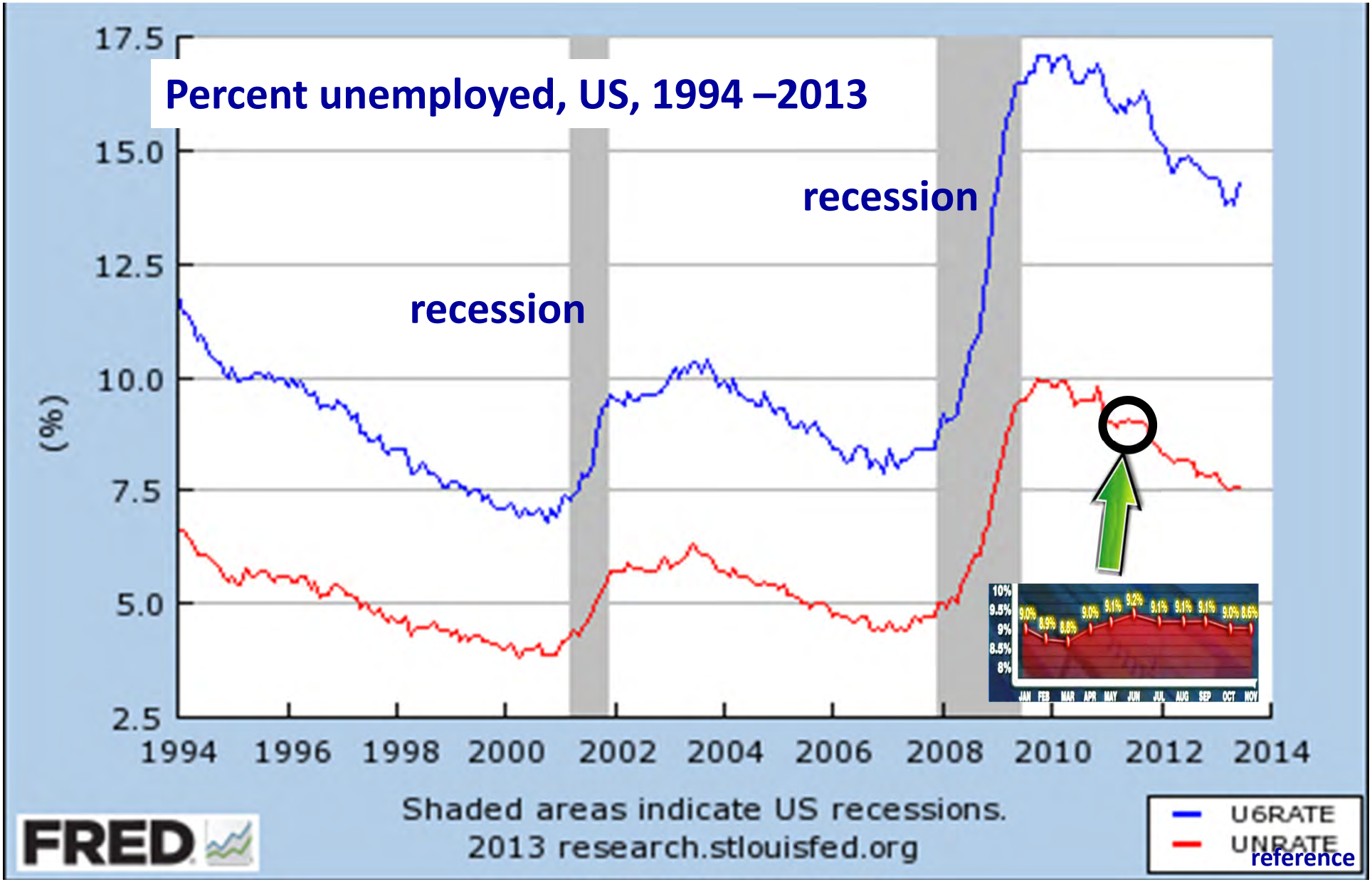
With the same data drawn truthfully, we see what appears to be a dramatic drop in unemployment…

Labor force status: Unemployment rate
Type of data: Percent or rate
Age: 16 years and over



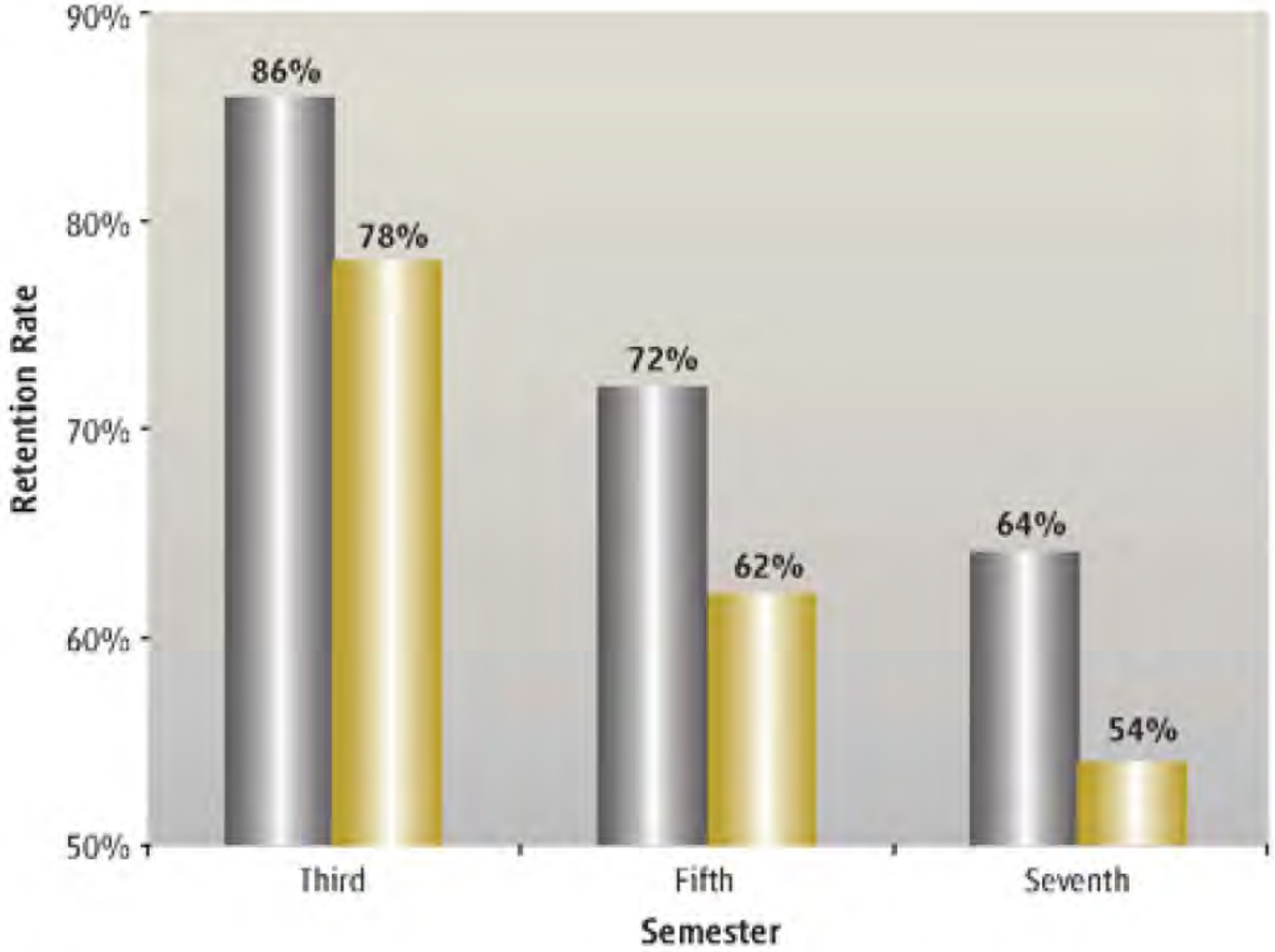
Bureau of Labor Statistics

… until we add more data for context.



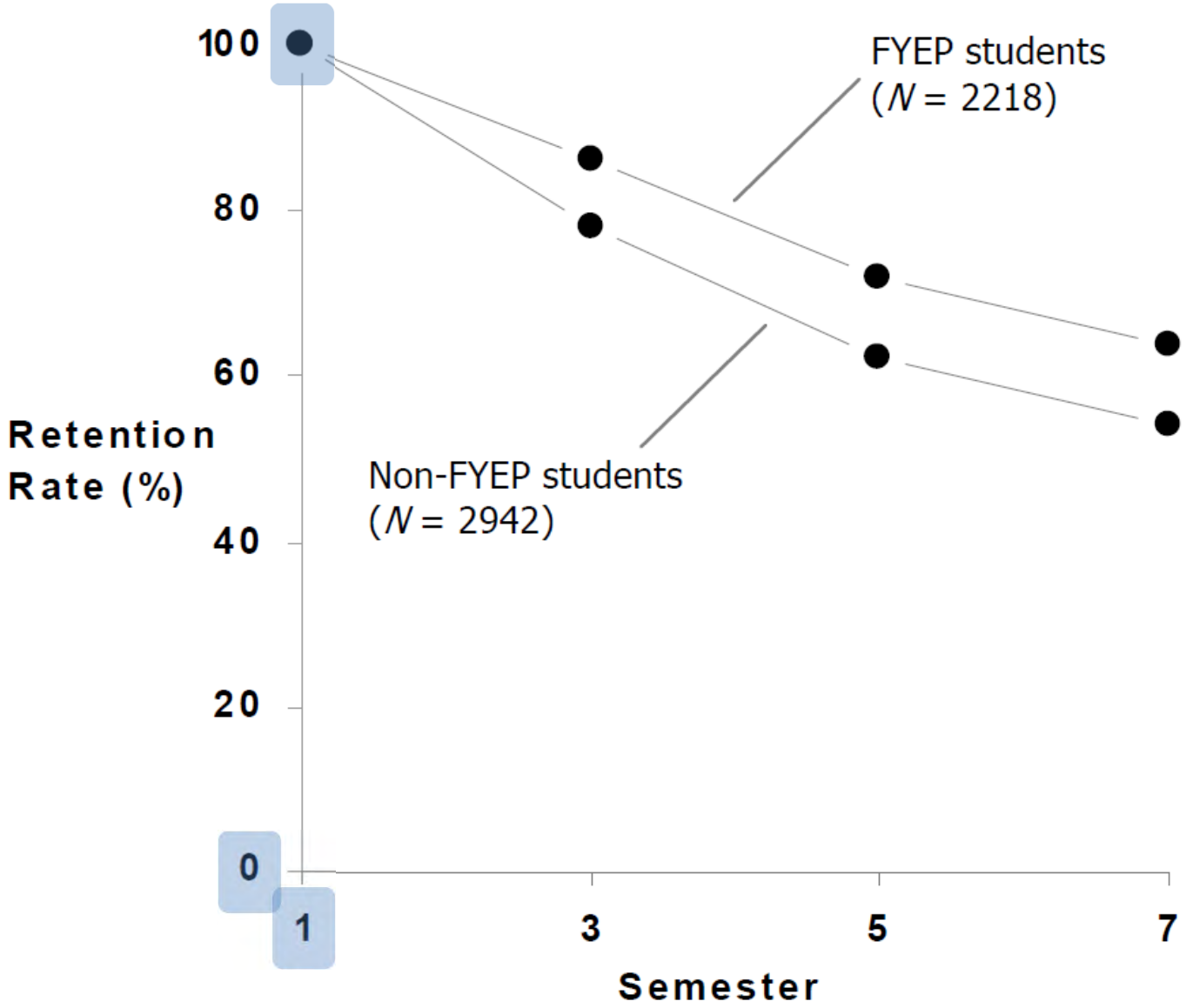
an inadvertent deception

bars that don't start at zero
missing data



Gains in retention. The FYEP course improved retention of engineering students into the third, fifth, and seventh semester. There were 2128 students who took the FYEP course (gray) and 2942 students who did not (gold). All retention gains over expected retention rates shown are significant ($P < 0.05$).

reference



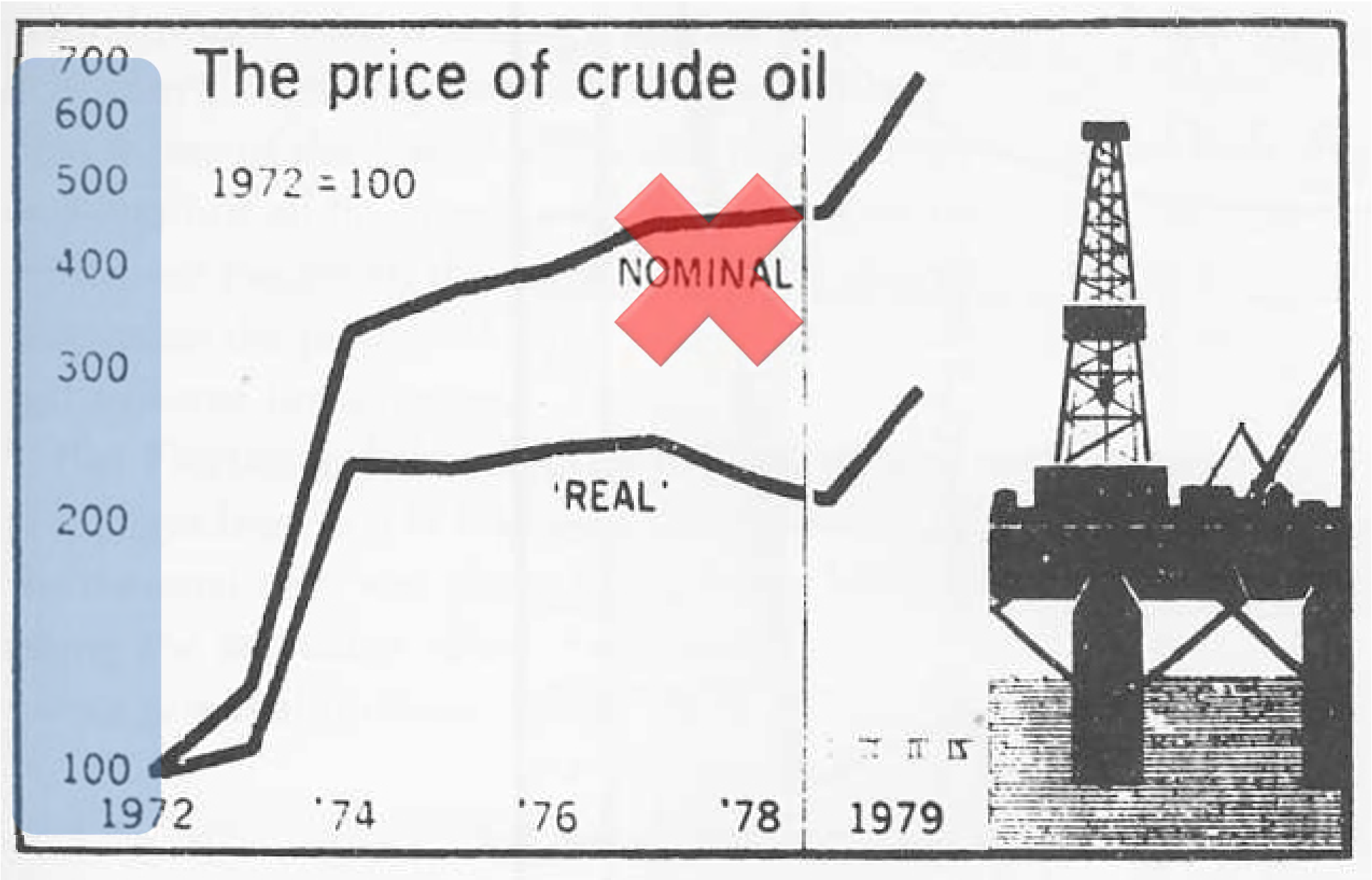
First-year gains in retention. The primary impact of the first-year engineering projects (FYEP) course is in the higher retention rate in the third semester. Subsequently, both groups lose students at about the same rate with a persistent 10% difference between FYEP and non-FYEP students.

reference

panic journalism

scale is misleading
shows non-constant dollars
lack of context

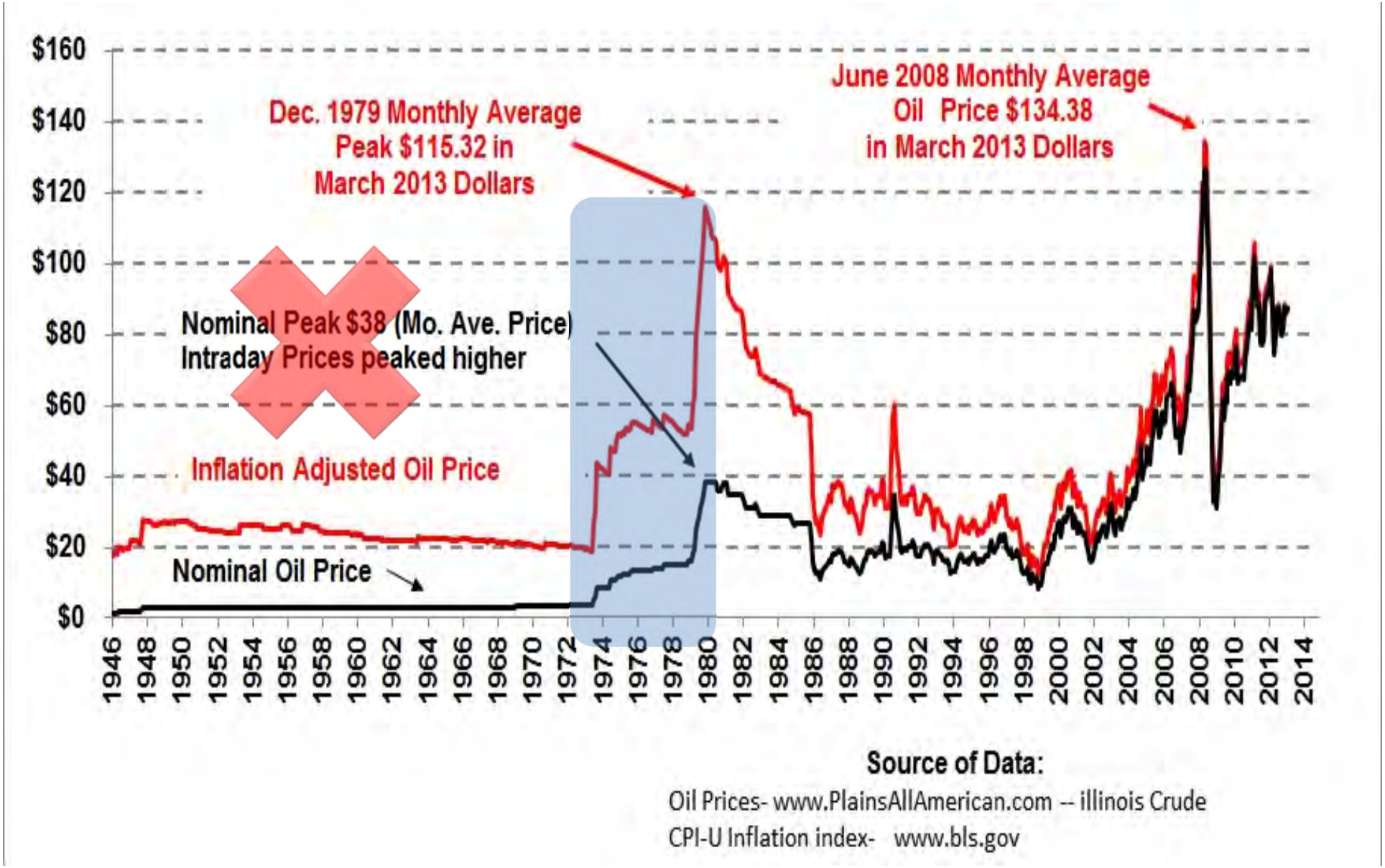
The price of crude oil is out of control.



Sunday Times (12/16/79) via Tufte

reference

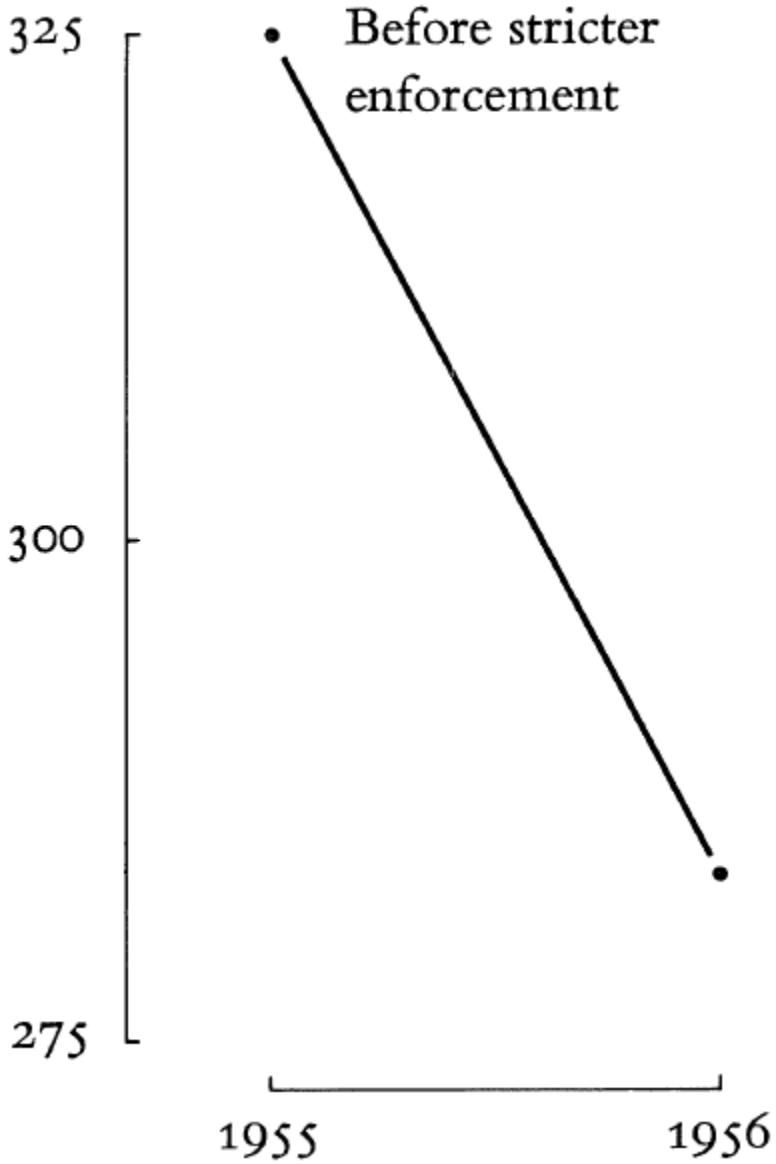
Monthly crude oil prices, constant 2013 dollars



reference

lie by omission

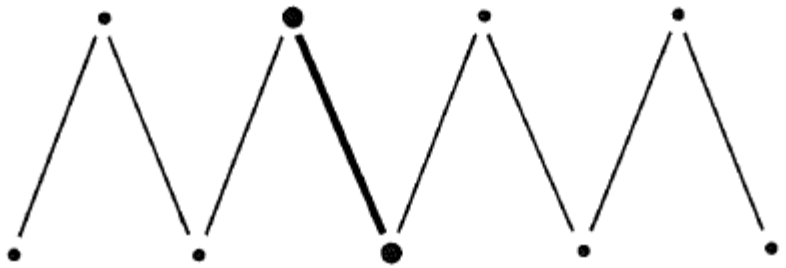
missing data
lack of context



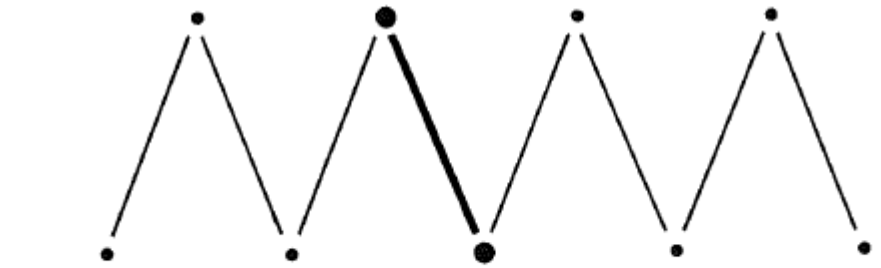
Connecticut Traffic Deaths,
Before (1955) and After (1956)
Stricter Enforcement by the Police
Against Cars Exceeding Speed limit

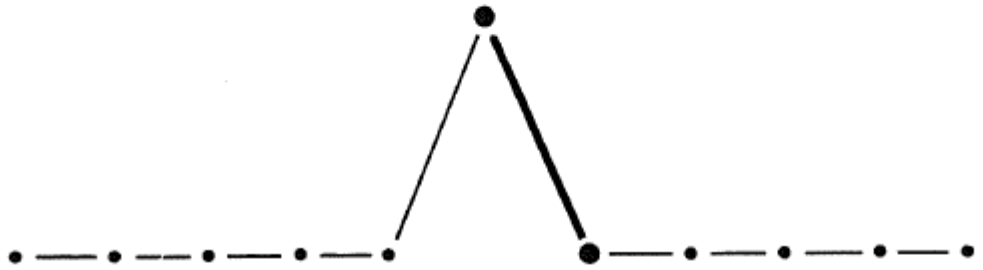
Compared to what?

reference

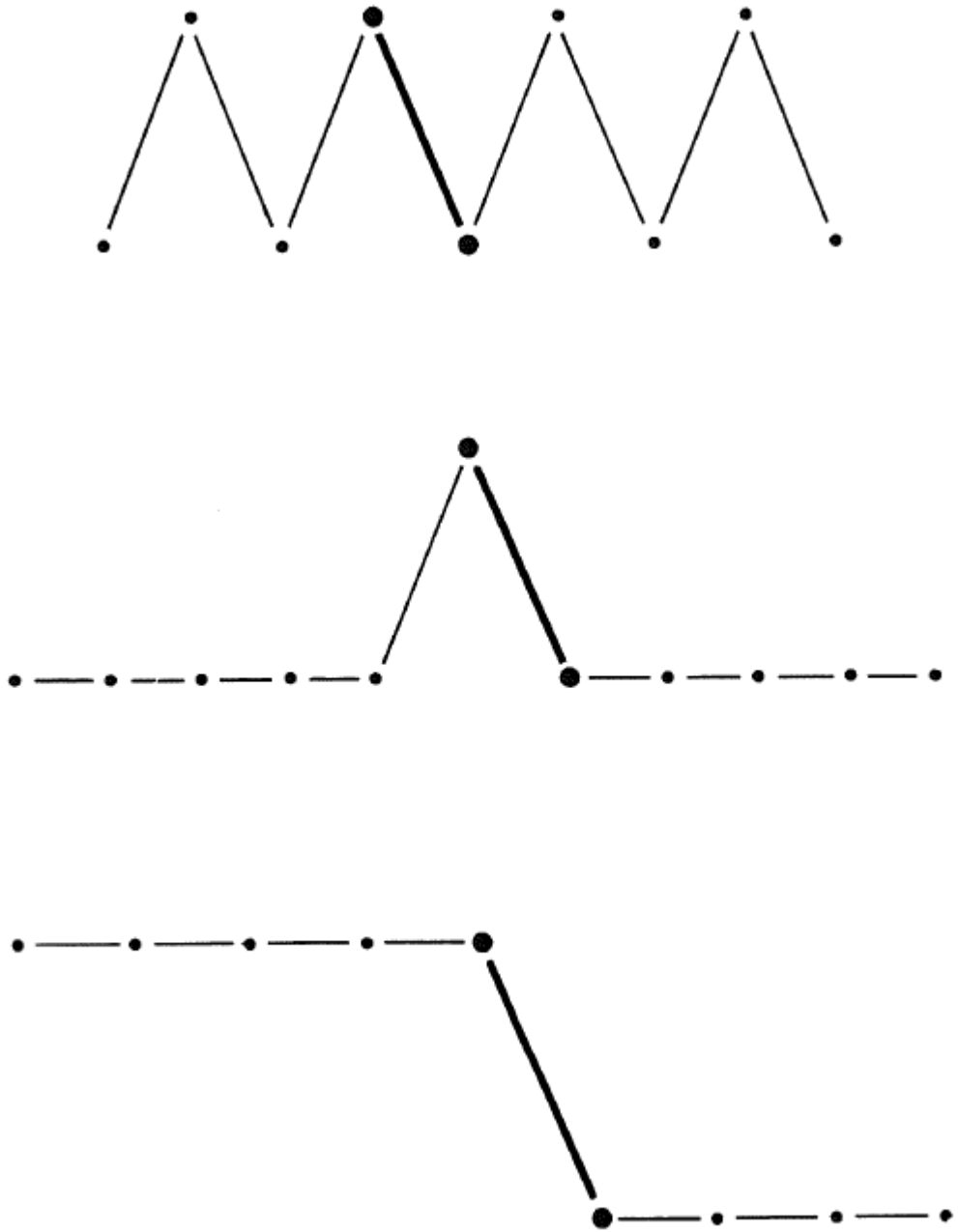


reference

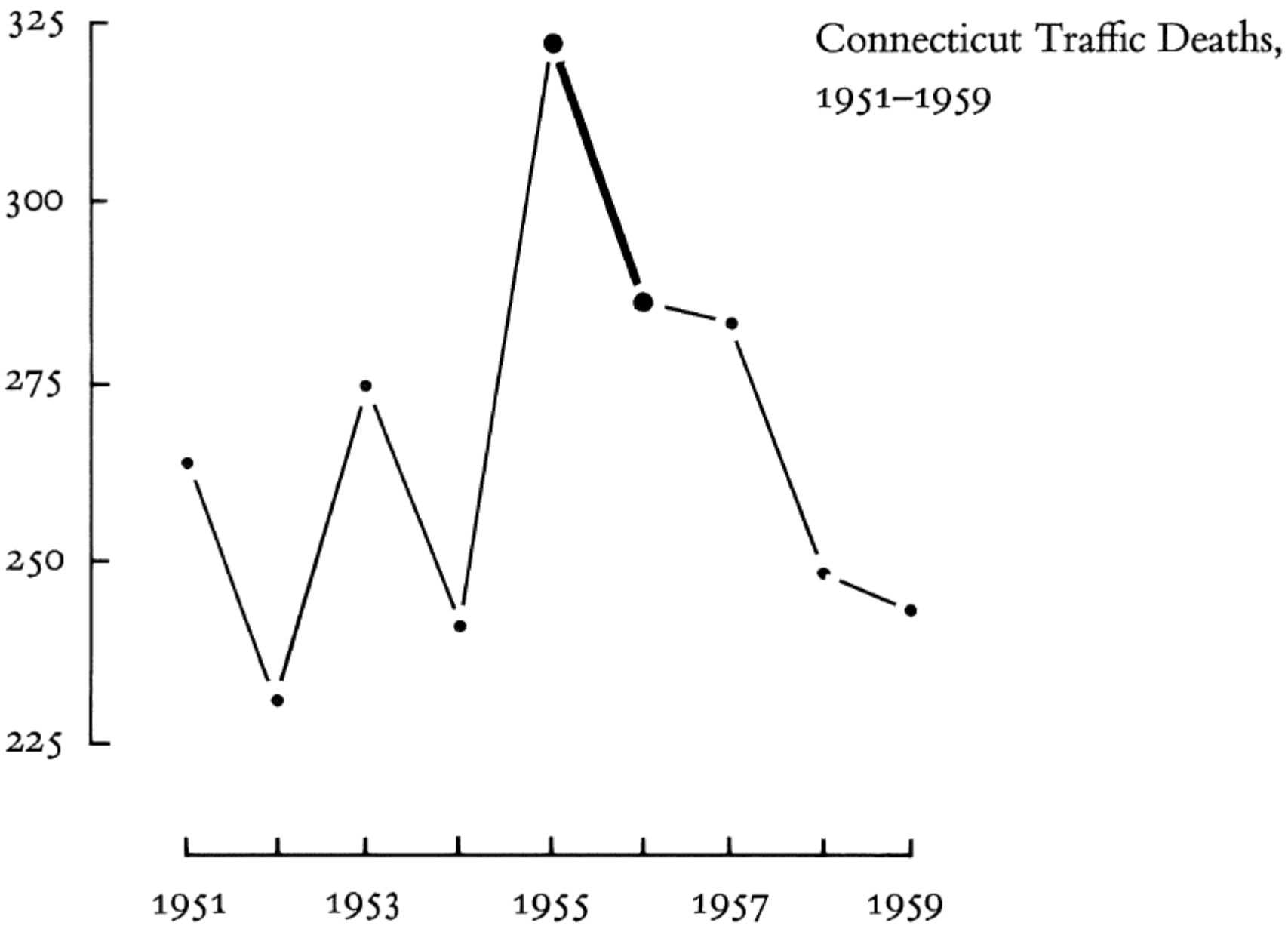




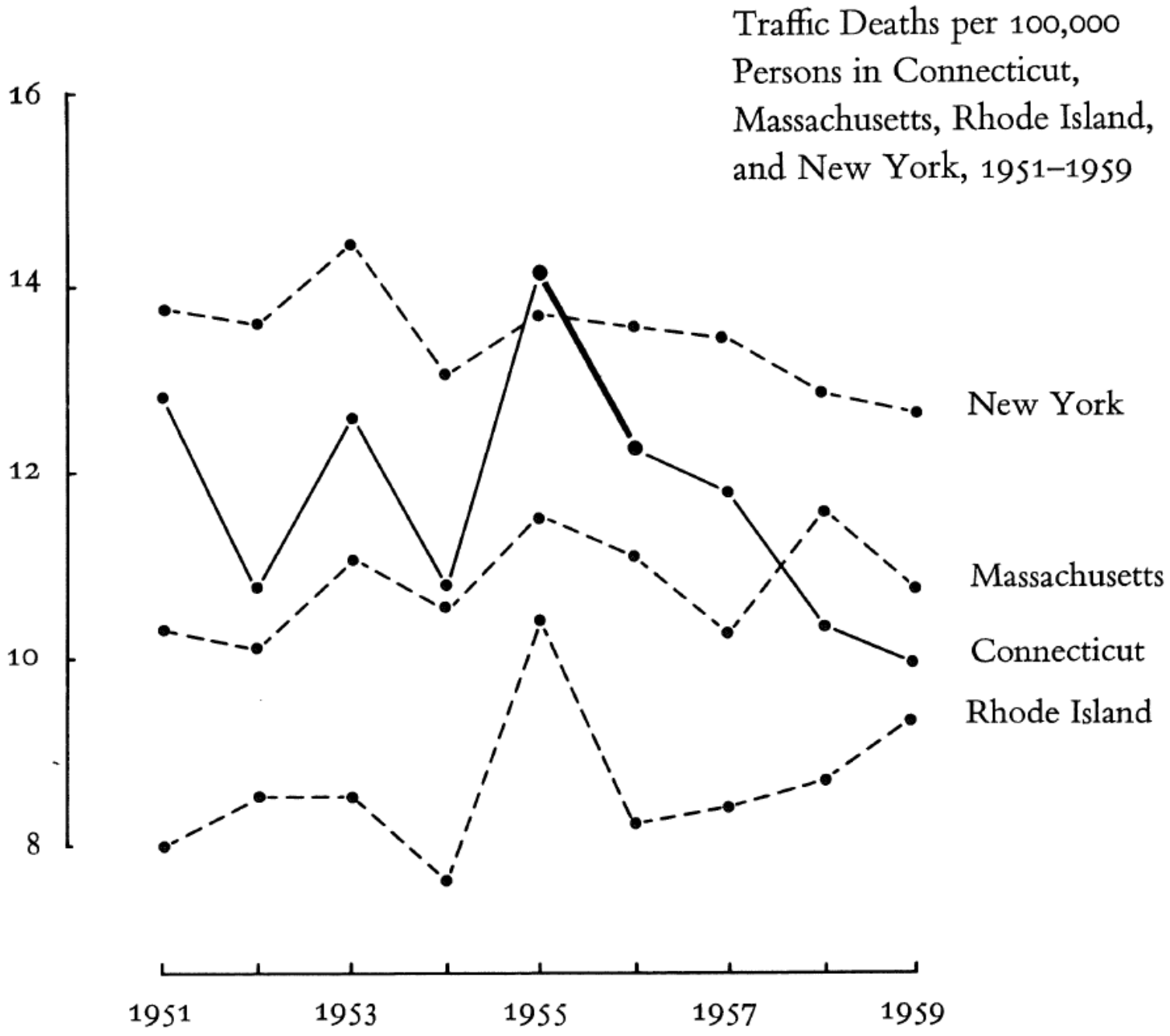
reference



reference


Connecticut Traffic Deaths,
1951–1959
325
300
275
250
225
1951
1953
1955
1957
1959

reference


Traffic Deaths per 100,000
Persons in Connecticut,
Massachusetts, Rhode Island,
and New York, 1951–1959
16
14
12
10
8
1951
1953
1955
1957
1959
New York
Massachusetts
Connecticut
Rhode Island

reference

lie by emphasizing the trivial

A graph from a recent paper on the progress of engineering students through a math sequence.

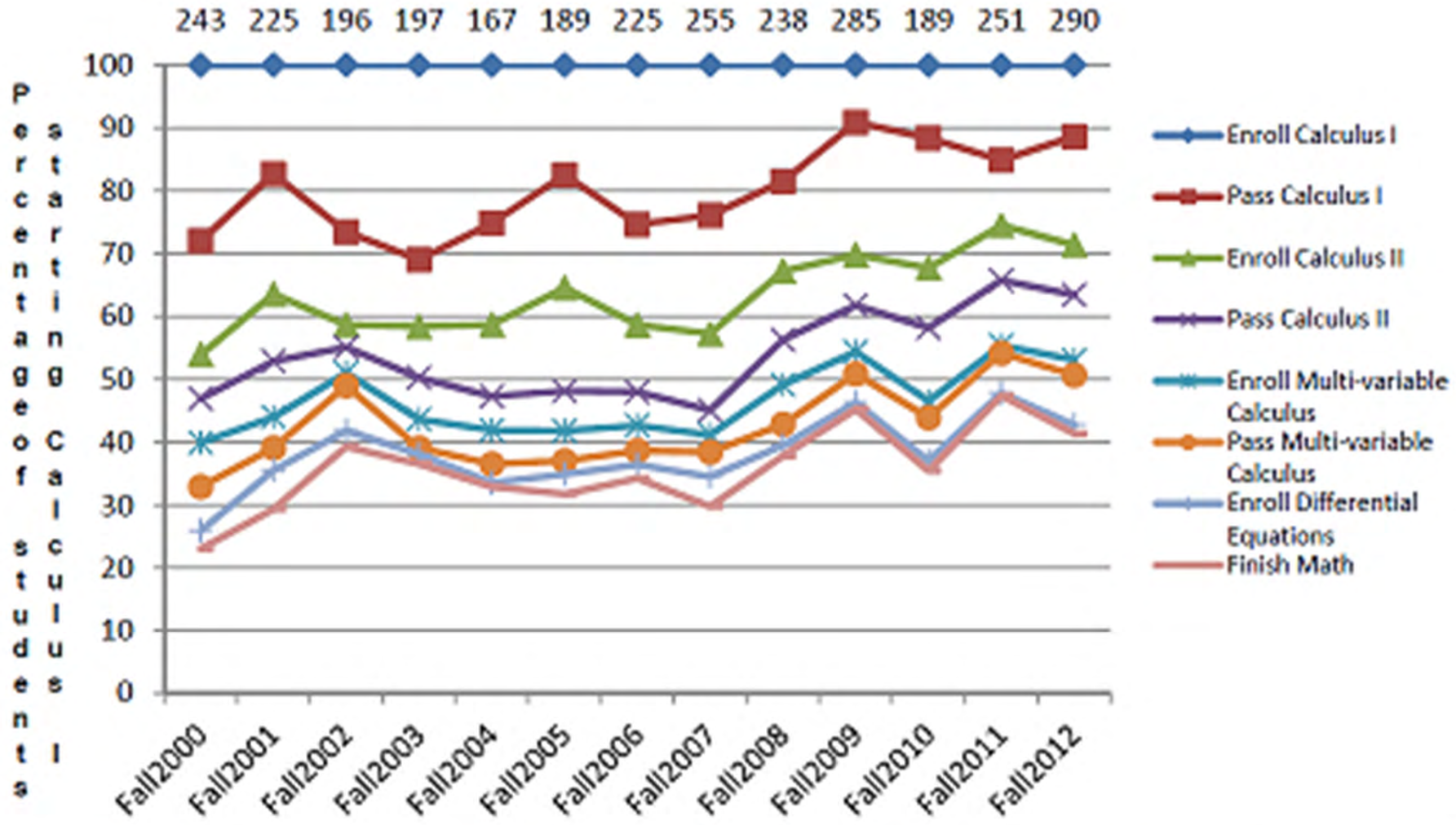
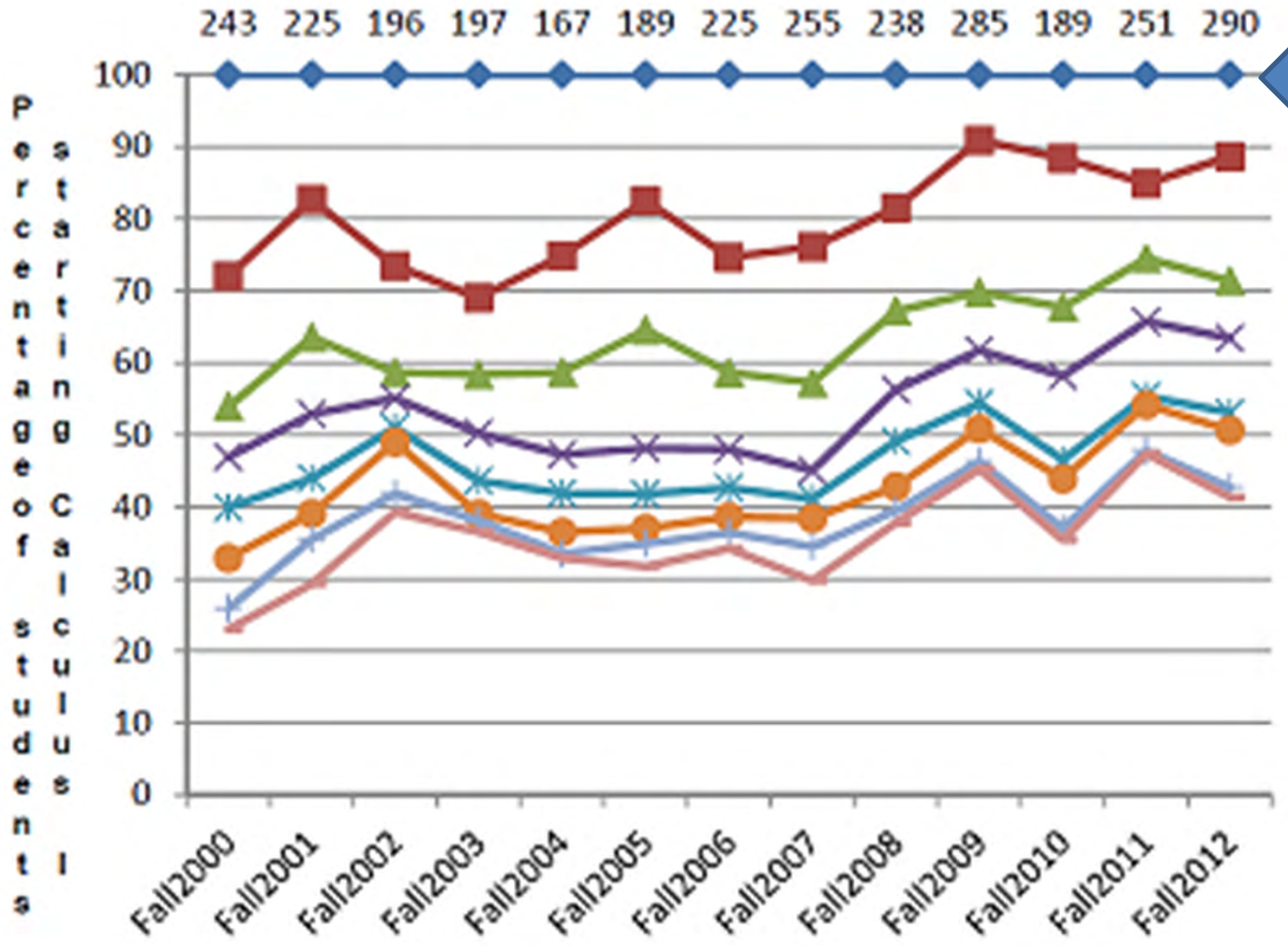


Figure 4. Female engineering student progress through the four-course mathematics sequence in consecutive semesters.

reference

The independent variable is time.



enroll in Calc 1

Figure 4. Female engineering student progress through the four-course ma
consecutive semesters.

reference

The independent variable is time, but the story lies in the differences between the lines.

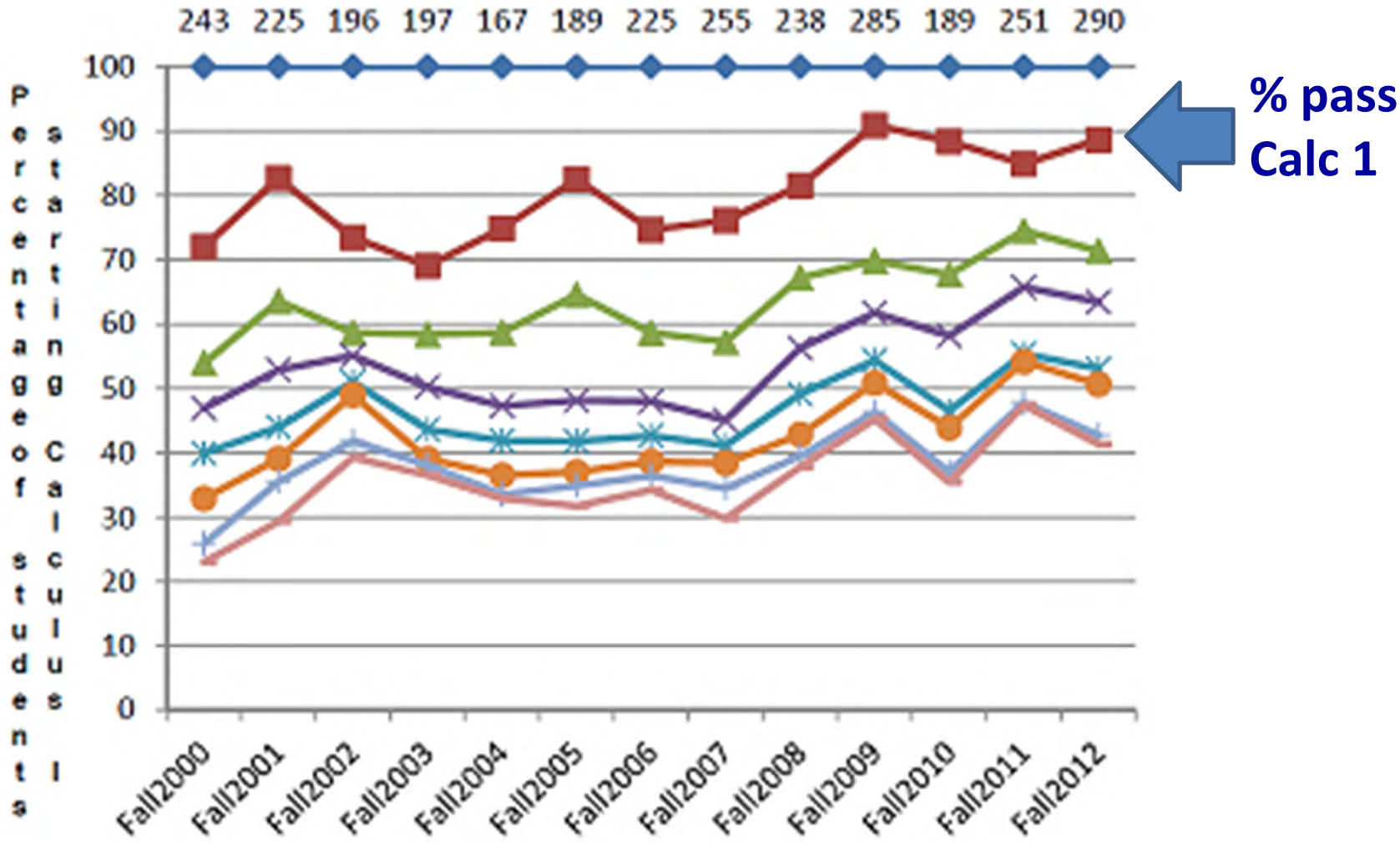
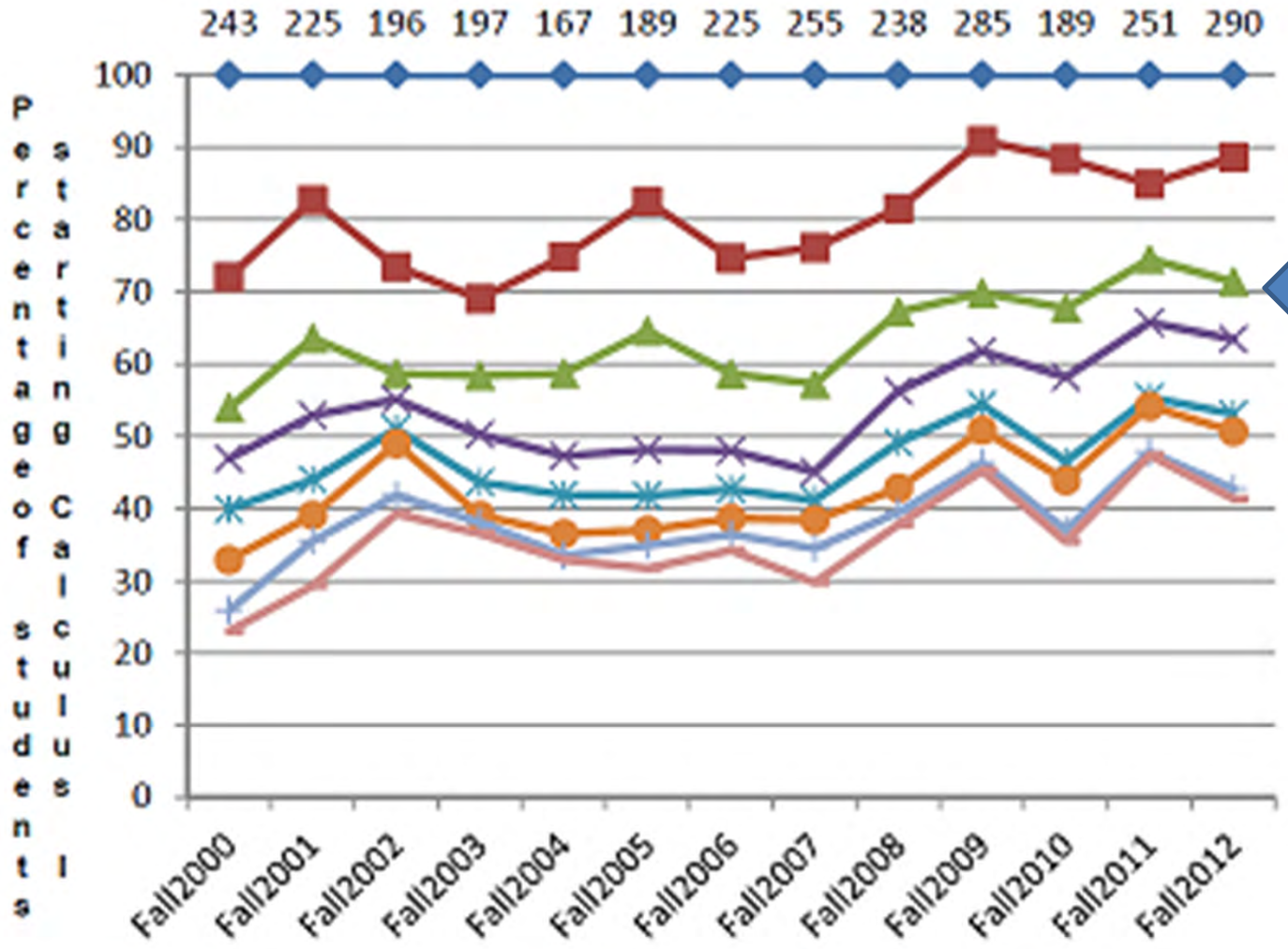


Figure 4. Female engineering student progress through the four-course ma
consecutive semesters.

reference

The independent variable is time, but the story lies in the differences between the lines.

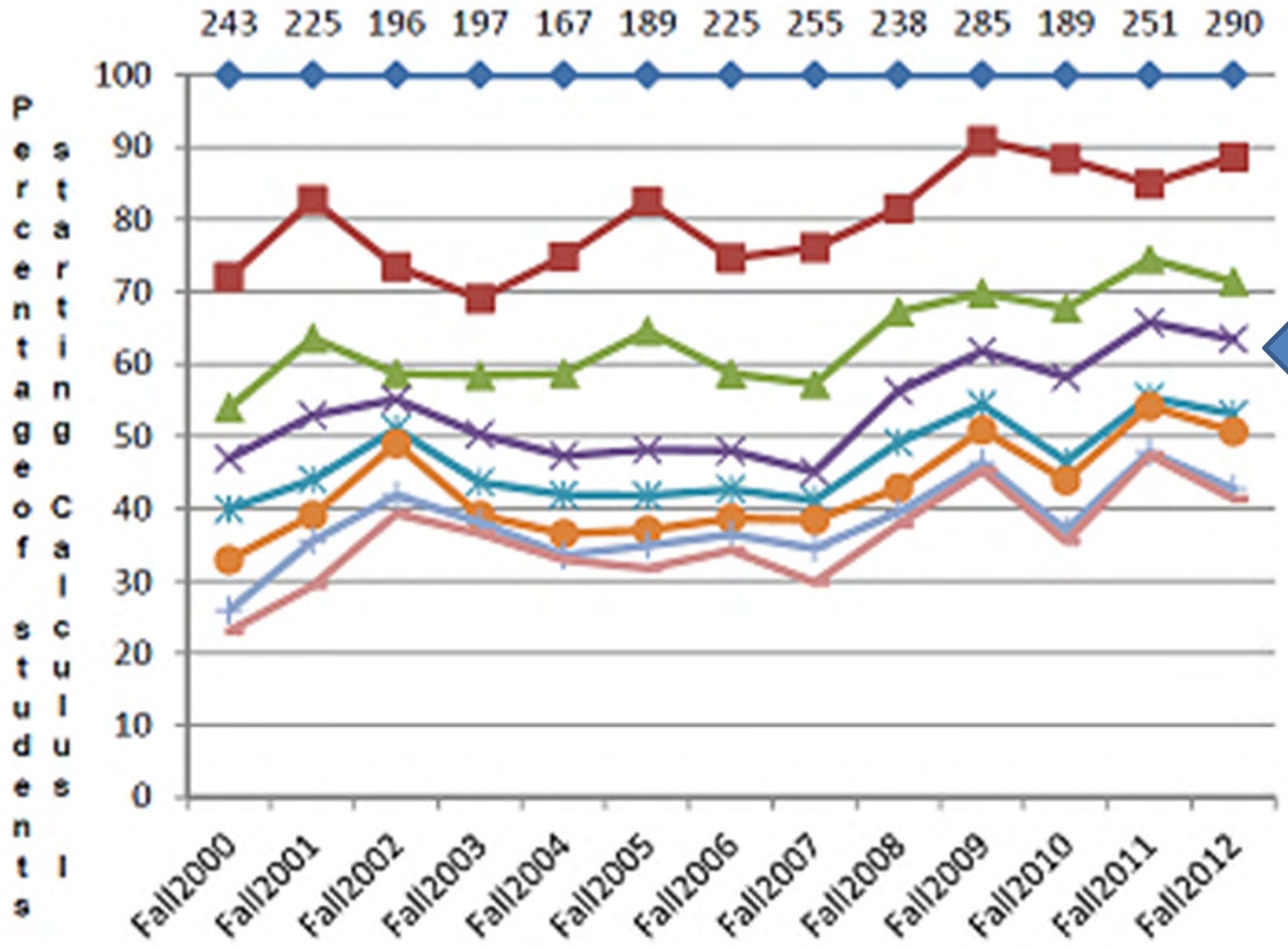


% enroll Calc 2

Figure 4. Female engineering student progress through the four-course ma
consecutive semesters.

reference

The independent variable is time, but the story lies in the differences between the lines.



% pass Calc 2

Figure 4. Female engineering student progress through the four-course ma
consecutive semesters.

reference

The independent variable is time, but the story lies in the differences between the lines.

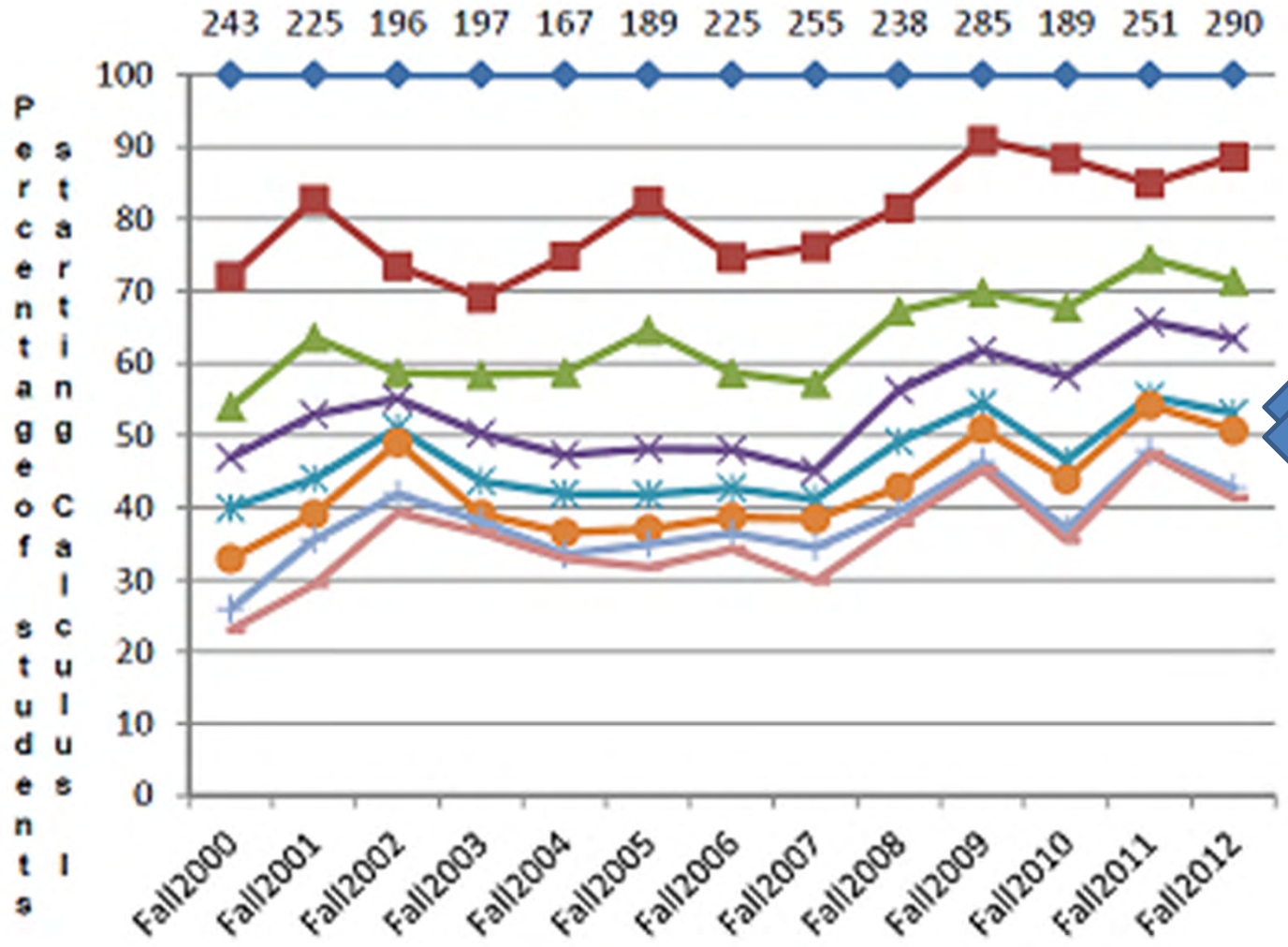
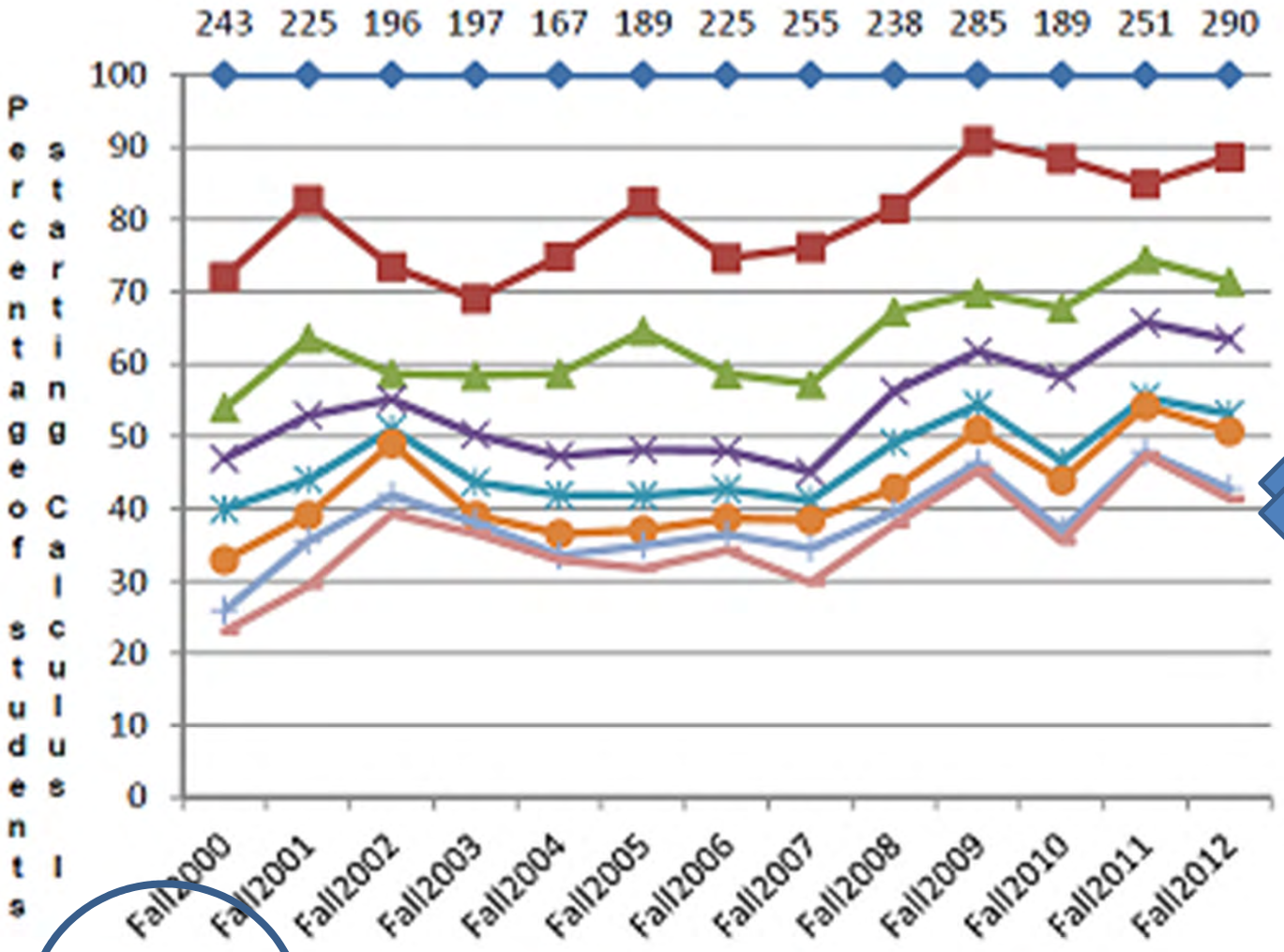


Figure 4. Female engineering student progress through the four-course ma
consecutive semesters.

reference

The independent variable is time, but the story lies in the differences between the lines … and gender.



% enroll DEs

% finish math seq

Figure 4. Female engineering student progress through the four-course ma
consecutive semesters.

reference

The same graph design for male students.

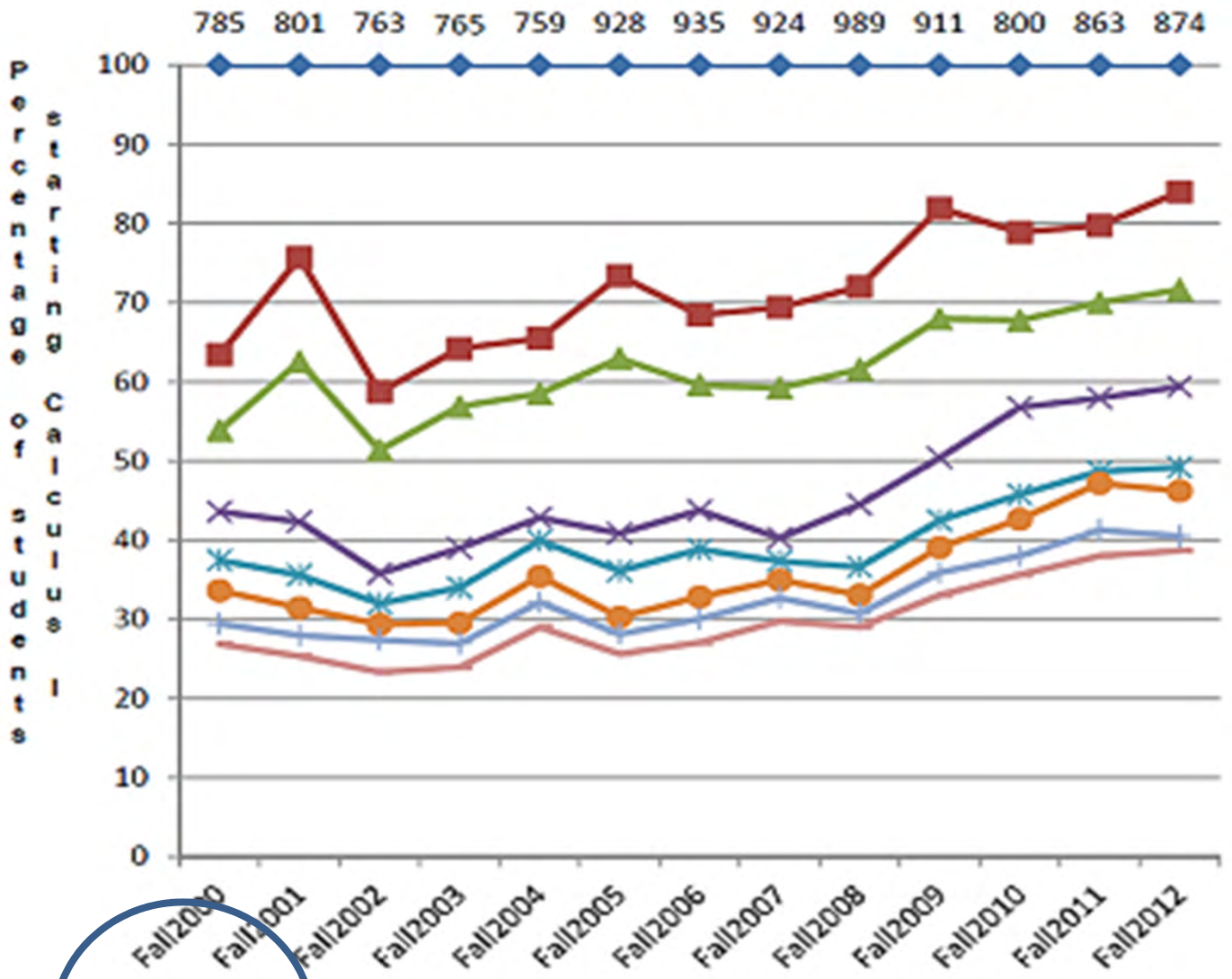
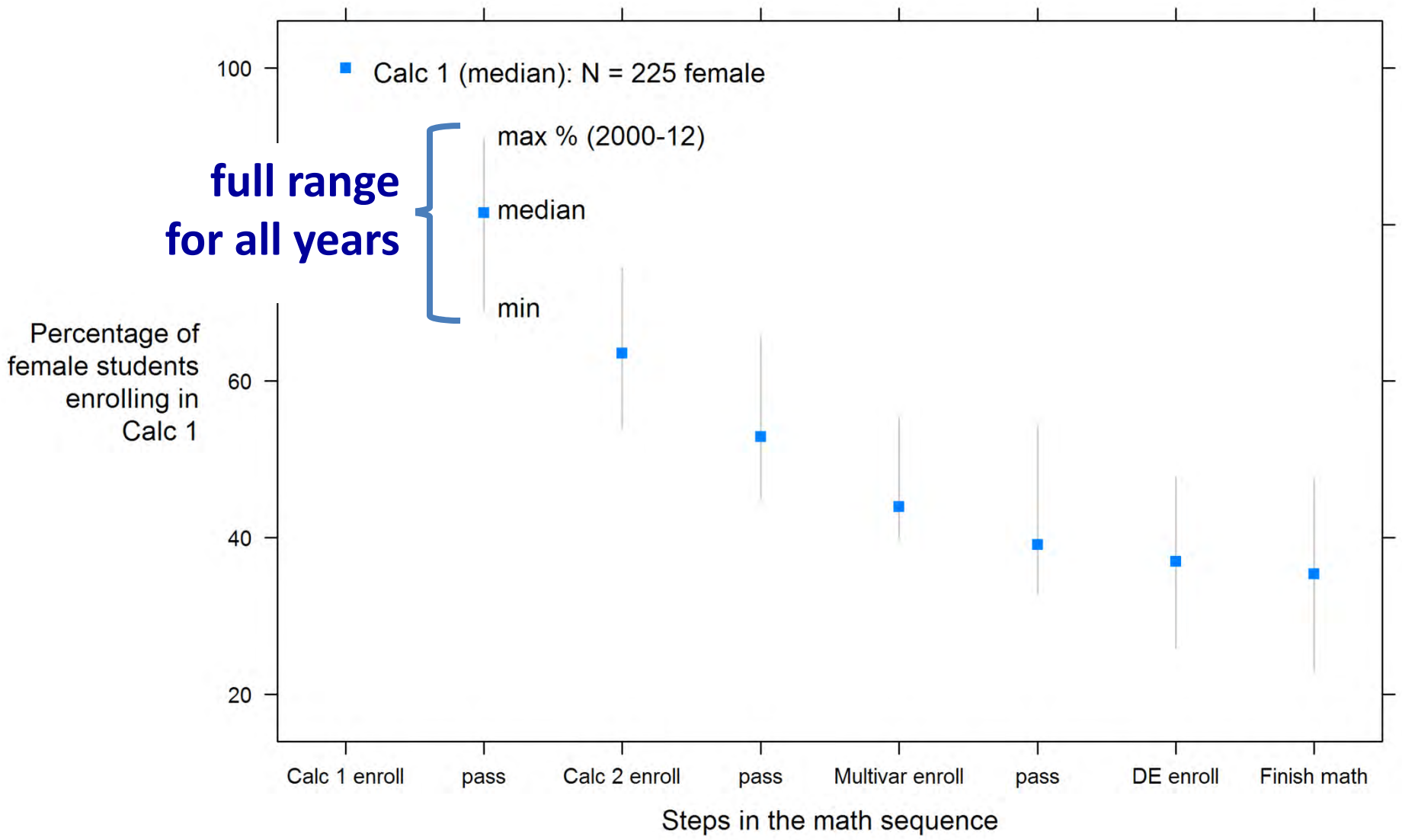


Figure 3. Male engineering student progress through the four-course mathe
consecutive semesters.

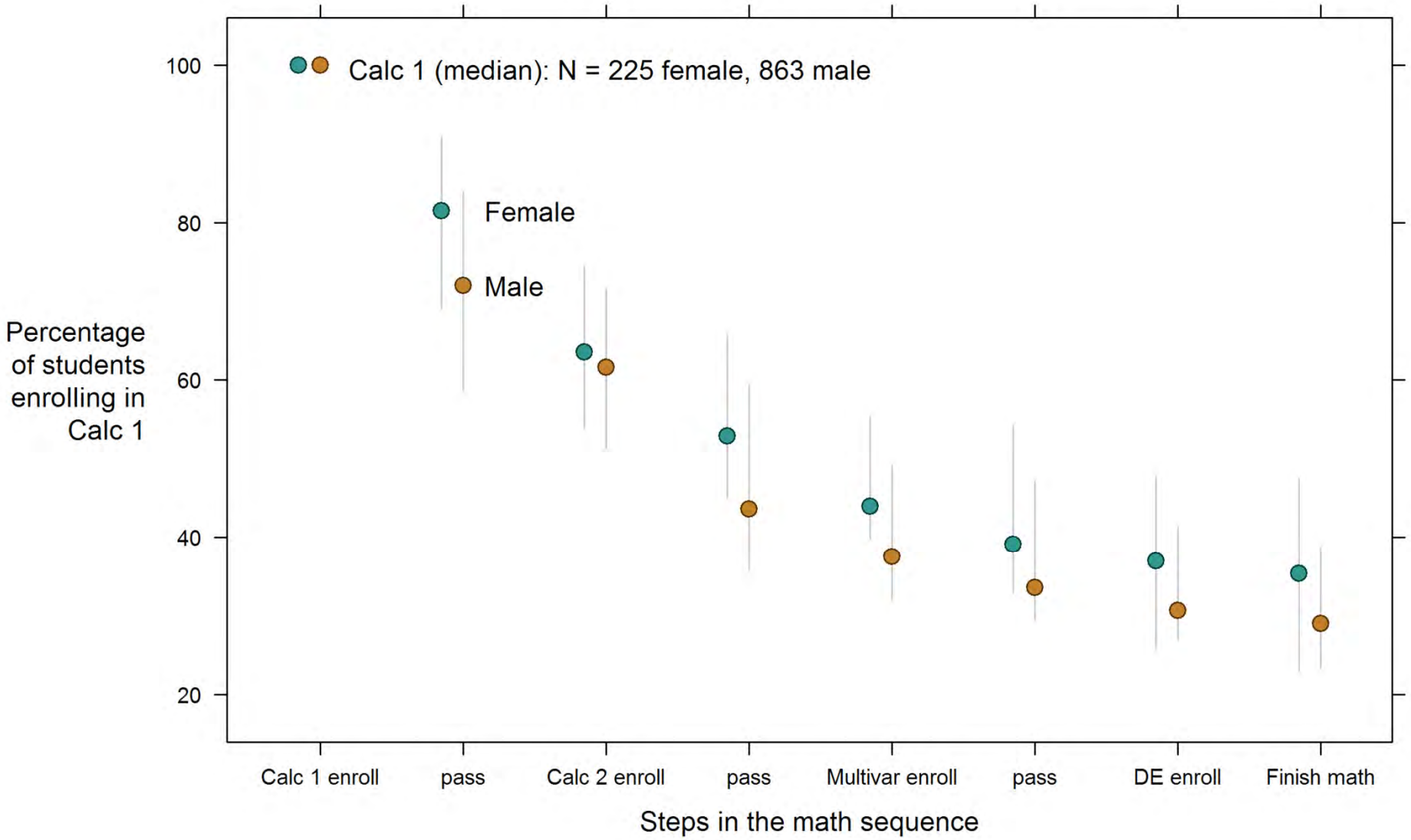
Comparing women & men not easy with this design.

reference

Steps in the math sequence are the independent variable (not time).



Dependencies are now directly observable.



The lie?
Concealing the story by emphasizing the trivial.

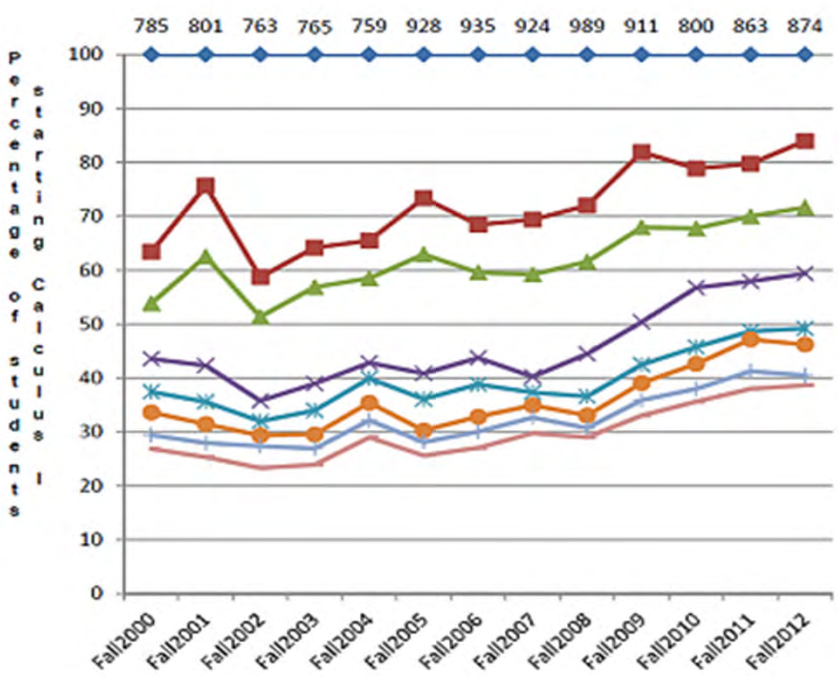


Figure 3. Male engineering student progress through the four-course mathe consecutive semesters.

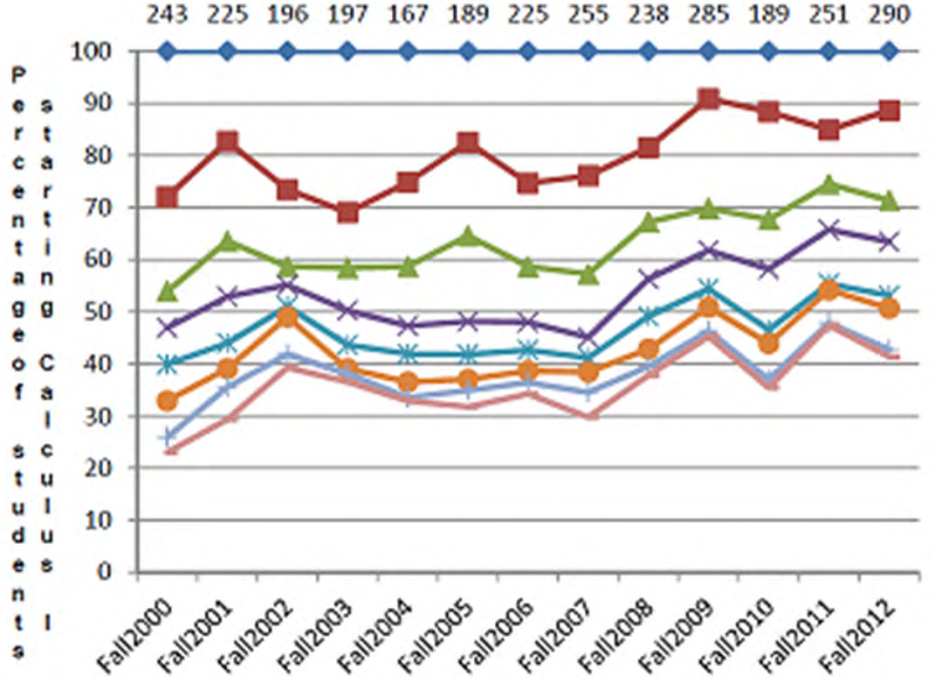


Figure 4. Female engineering student progress through the four-course ma consecutive semesters.

reference

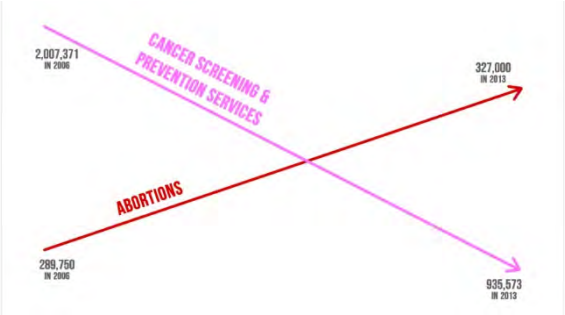
Edward Tufte

The main defense of the lying graphic is … "Well, at least it was approximately correct, we were just trying to show the general direction of change."

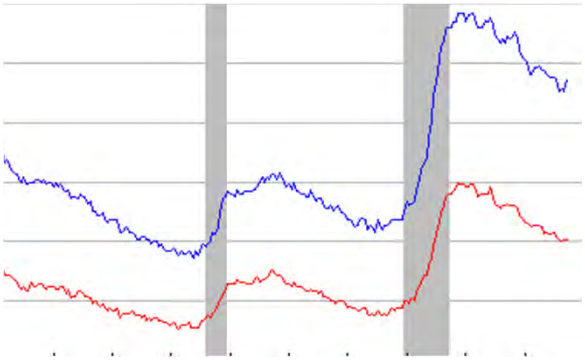
A second defense is, "although the design lies, the numbers are printed, as if not lying in one place justifies lying elsewhere."

reference

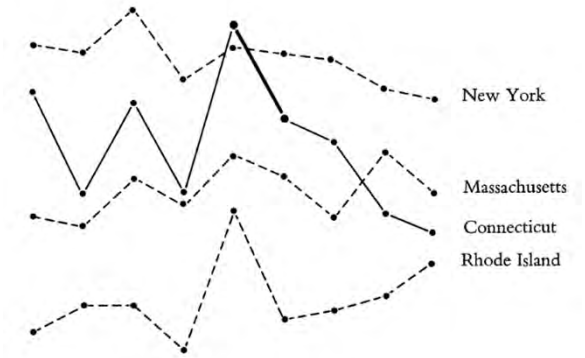
Implications for the designer.



Some graph designers abuse their power. Know the tricks used to deceive.



Avoid inadvertent deception due to laziness, bad design, and lack of context.



Don’t defend the lying graphic. Fix it!