Ronald Reagan was a statistician and other examples of learning from diverse sources of information

Andrew Gelman
Department of Statistics, Columbia University, New York

21 Mar 2013
“Are you better off than you were four years ago?”
“But the simplest way for each of you to judge recovery lies in the plain facts of your own individual situation. Are you better off than you were last year?”
Ronald Reagan was a statistician
Ronald Reagan was a statistician
Who do you know?

- How many people do you know named Nicole?
- How many people do you know named Anthony?
- How many lawyers do you know?
- How many people do you know who were robbed in the past year?
On average, you knew 0.6 Nickels

0.13% of Americans are named Nicole

Assume 0.13% of your acquaintances are Nickels

Estimate: on average, you know \( \frac{0.6}{0.0013} = 450 \) people

On average, you know 0.8 Anthonys

0.31% of Americans are named Anthony

Estimate: on average, you know \( \frac{0.8}{0.0031} = 260 \) people

Why do these differ?
Perception vs. reality: Names

Ronald Reagan was a statistician
Social Structure!

How many Jacquelines do you know?

How many men in prison do you know?

Andrew Gelman
Ronald Reagan was a statistician
The Gay Penumbra

Marriage Support: Everyone

Marriage Support: Democrats

Marriage Support: Republicans

Marriage Support: Independents

Andrew Gelman

Ronald Reagan was a statistician
How did this happen?

United States of Canada

Jesusland
It makes no sense!

Andrew Gelman
Ronald Reagan was a statistician
It’s no illusion

Republican vote by state in 2004

Vote share for George Bush
Average income within state

Andrew Gelman
Ronald Reagan was a statistician
What has changed?

1976 election
Average income within state
Vote share for Gerald Ford
AL
AZ
AR
CA
CO
CT
DEFL
GA
HI
ID
IL
IN
IA
KS
KY... ORPA
RISC
SD
TN
TX
UT
VT
VA WA
WV
WI
WY
$15,000 $20,000
30% 50% 70%

Republican vote by state in 1976

Andrew Gelman
Ronald Reagan was a statistician
It’s not just the South

Republican vote in rich states minus Republican vote in poor states

-20% 0 20%

All states

Southern states

Non-Southern states

1940 1960 1980 2000

Andrew Gelman

Ronald Reagan was a statistician
The rich sill vote Republican

Bush vote in 2004 by income

Vote share for Bush
0 $100,000 $200,000

Individual income

Vote share for Bush
0 $100,000 $200,000
30% 50% 70%

Bush vote in 2004 by income

2006 House exit polls

Republican vote share

Northeast
Midwest
South
West

Andrew Gelman
Ronald Reagan was a statistician
Rich counties are more Republican

Andrew Gelman

Ronald Reagan was a statistician
But not in the West!
“Opiate of the masses” vs. “Postmaterialism”

- “I don’t know that atheists should be considered citizens, nor should they be considered patriots. This is one nation under God.” — George H. W. Bush
- “Very few of us . . . could name even five NASCAR drivers, although stock-car races are the best-attended sporting events in the country.” — David Brooks
- Two theories:
  - *Opiate of the masses*: Rich people vote their interests, poor people vote “Gods, guns, and gays”
  - *Postmaterialism*: Poor people vote based on economics, rich people have the luxury to vote on social issues
“Opiate” no, “Postmaterialism” yes

Bush vote in 2004 by income and religious attendance

Probability of voting for Bush

poor middle-income rich
40% 50% 60% 70%

if you attend church more than once/week

if you attend once or twice/month

if you never attend church

Andrew Gelman

Ronald Reagan was a statistician
Similar patterns in red and blue America

Republican States

Battleground States

Democratic States

Andrew Gelman

Ronald Reagan was a statistician
2000: Do you support school vouchers?

<table>
<thead>
<tr>
<th>Income under $20,000</th>
<th>$20–40,000</th>
<th>$40–75,000</th>
<th>$75–150,000</th>
<th>Over $150,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>All voters</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>White Catholics</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>White evangelical Protestants</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>White non-evang. Protestants</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>White other/ no religion</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>Blacks</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>Hispanics</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
<tr>
<td>Other races</td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
<td><img src="..." alt="Maps" /></td>
</tr>
</tbody>
</table>

The state is left blank where a category represents less than 1% of the voters of a state.
Data don’t always “speak for themselves”
Arthur “not David” Brooks in the New York Times:
“People at the extremes are happier than political moderates. . . . none, it seems, are happier than the Tea Partiers . . .”

Jay Livingston (sociology, Montclair State University) looks up the data in the General Social Survey . . .
“None, it seems, are happier than the Tea Partiers …”

Chart for YEAR = 4 (2009-2010)

**GENERAL HAPPINESS BY THINK OF SELF AS LIBERAL OR CONSERVATIVE**

<table>
<thead>
<tr>
<th>Percent</th>
<th>EXTREMELY LIBERAL</th>
<th>LIBERAL</th>
<th>SLIGHTLY LIBERAL</th>
<th>MODERATE</th>
<th>SLIGHTLY CONSERVATIVE</th>
<th>CONSERVATIVE</th>
<th>EXTREMELY CONSERVATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERY HAPPY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRETTY HAPPY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOT TOO HAPPY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
General Happiness by Think of Self as Liberal or Conservative

- Extremely Liberal
- Liberal
- Slightly Liberal
- Moderate
- Slightly Conservative
- Conservative
- Extremely Conservative

Legend:
- Very Happy
- Pretty Happy
- Not Too Happy
Local Data, National Coordination, Local Decisions

Older Model, left: Observed indoor radon levels, based on short-term measurements collected by the E.P.A. and most states. The measurements, taken in basements in the winter, may overestimate true exposure. New Model, right: The predicted average long-term levels of radon that people may actually breathe in their homes.

The New York Times
A data party!
A big data party!
Big data, little data,
red data, blue data,
yellow data, green data,
black data, and white data,
are all at a data party!
What a data party!