



Polls, surveys and rankings

Three wrongs and some rights

MATT CARMICHAEL, PETER GADE, CHRIS JACKSON

GAME CHANGERS



Contents

03	Preface	26	Interpretation Error
07	Introduction	33	Examples of cool stuff
08	Sampling Error	39	Contacts
17	Measurement Error		

Who will win the presidency?



Chance of winning

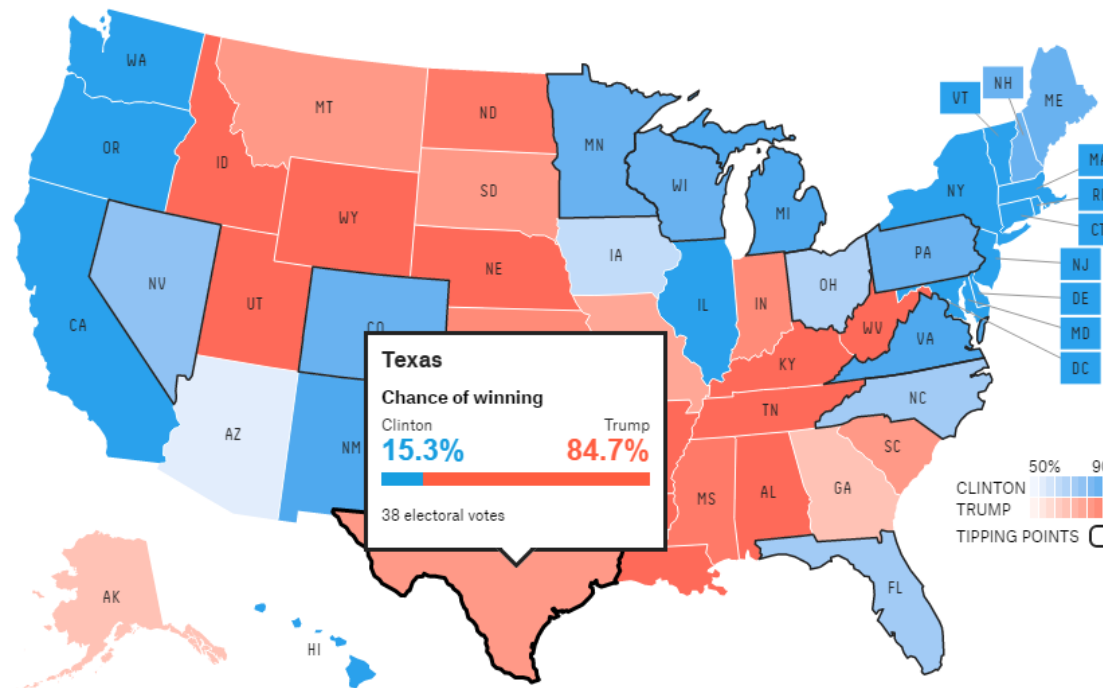


Hillary Clinton

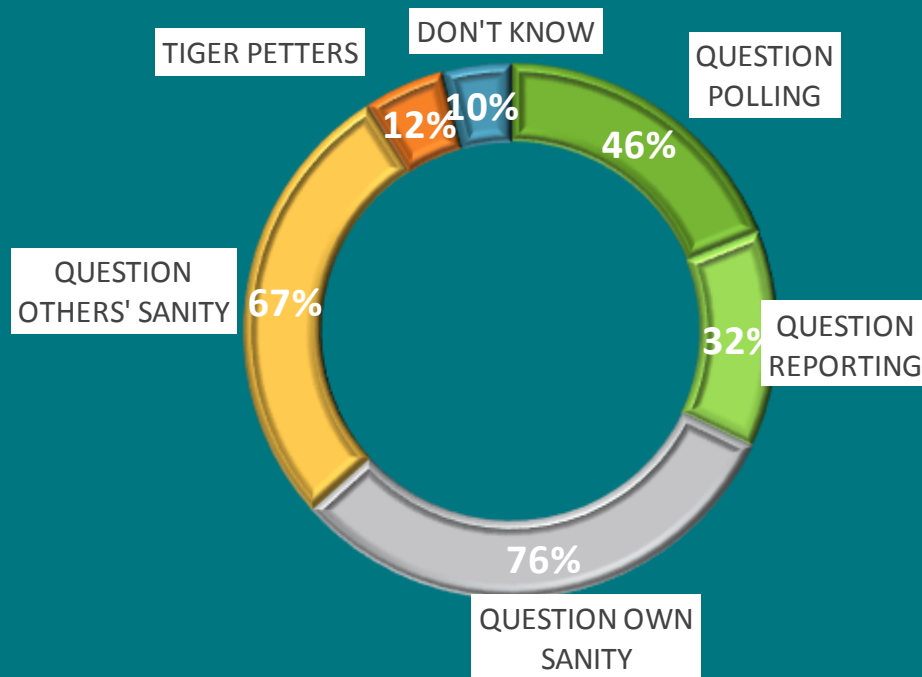
87.1%

Donald Trump

12.9%



THE ELECTION MADE AMERICANS QUESTION SANITY OF THEMSELVES/OTHERS, ACCORDING TO MATT'S SURVEY





Donald J. Trump ✓

@realDonaldTrump



Follow



Any negative polls are fake news, just like the CNN, ABC, NBC polls in the election. Sorry, people want border security and extreme vetting.

RETWEETS

27,897

LIKES

118,258





Donald J. Trump  @realDonaldTrump · Feb 8

'Trump administration seen as more truthful than news media'
bit.ly/2k4b0imEmerson...



40K



18K



89K

Who we are

PETER GADE, PHD

CURRENT: GAYLORD CHAIR OF JOURNALISM AT THE UNIVERSITY OF OKLAHOMA

PAST: CO-EDITOR OF *CHANGING THE NEWS* (2011) AND CO-AUTHOR OF *TWILIGHT OF PRESS FREEDOM* (2001)

CHRIS JACKSON

CURRENT: VICE PRESIDENT, IPSOS PUBLIC AFFAIRS

PAST: DIRECTOR, KRC RESEARCH; RESEARCHER, BELDEN RUSSONELLO AND STEWART;

MATT CARMICHAEL

CURRENT: DIRECTOR OF EDITORIAL STRATEGY FOR IPSOS NORTH AMERICA

PAST: EDITOR, RESEARCHER AT: ADVERTISING AGE; CRAIN'S CHICAGO BUSINESS; LIVABILITY.COM

The first wrong: **SAMPLING ERROR**

Who are the people in your survey and why does that matter?

Types of polls and sampling

How the sample is drawn is essential to determining a poll's validity or truthfulness.

Sampling error occurs through the process of determining whom to include as participants in the poll.

To make inferences from a sample to a population – to generalize, you want a sample that is representative of the population.

To control for sampling bias, researchers allow the “process of pure chance” (randomness) to work.

Random Samples

Randomness means totally left to chance.

To ensure randomness:

- 1) Sample selected in a way only chance operates, and
- 2) every member of the population has an equal chance of being chosen.

Non-random samples (convenience, volunteer, purposive, quota) are unlikely to produce data that can be generalized to populations.

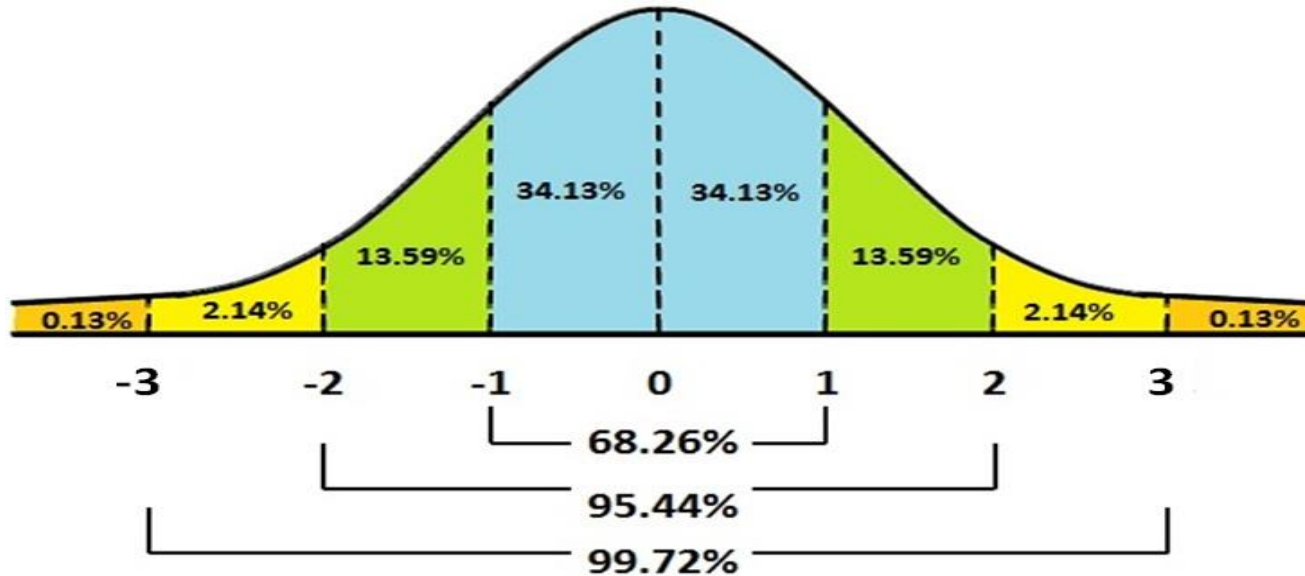
Polls and sampling error

When researchers refer to *margin of error*, they are referring to *sampling error (SE)*.

Sampling error (or standard error) is a measure of how much a sample varies from other samples we could draw from the same population.

Put another way, if we drew additional samples of the same size from the same population, how similar/different would they be?

The normal curve



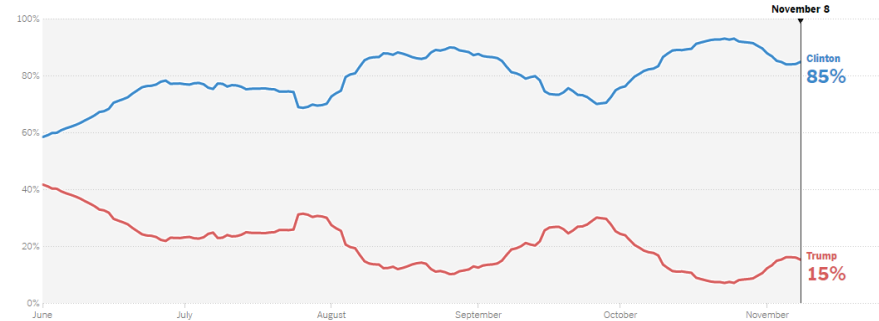
The **Central Limit Theorem** (normal curve) indicates that 95% of additional theoretical samples would fall within ± 2 SE of the percentages in the sample.

New York Times on election day

CLINTON: 45.9%

TRUMP: 42.8%

- The percentages were a national average of leading polls.
- Sampling methods differed across polls. Data were weighted.
- No margin of error was reported.



Interpreting the poll

LET'S SUPPOSE THE *NYT* ELECTION DAY POLL RESULTED FROM A SINGLE NATIONAL RANDOM SAMPLE THAT PRODUCED A MARGIN OF ERROR (SE) OF +/- 3%.

Considering the margin of error, another poll taken from same population is 95% likely to return these results of support for each candidate:

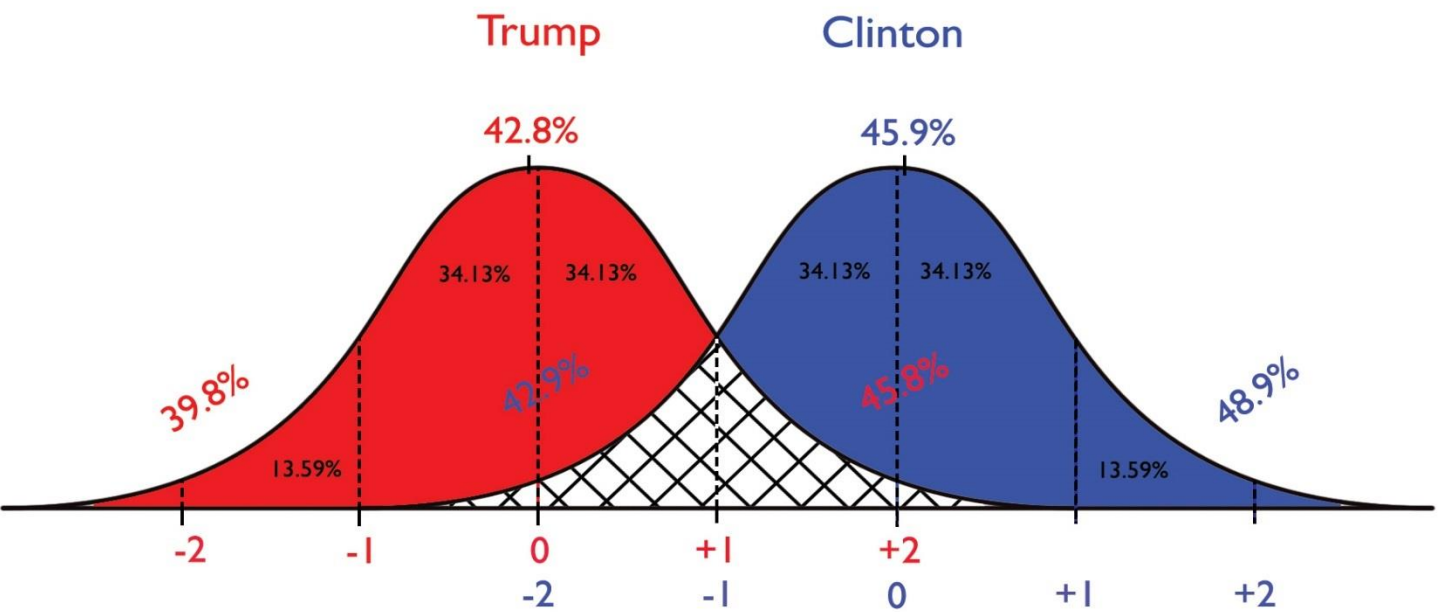
Clinton: $45.9\% \pm 3\% = 42.9 \text{ to } 48.9\%$

Trump: $42.8\% \pm 3\% = 39.8 \text{ to } 45.8\%$

We can conclude (at 95% confidence level):

--Clinton and Trump are within margin of error – no clear leader

Illustrating Margin of Error



What went wrong?

The poll was a national average of polls that could calculate sampling error (random) *and those that could not*. A mish-mash.

Weighting. This introduces the likelihood of additional sampling error.

Polls respondents, especially those online, are not representative of the broader population.

*These factors contributed to **sampling error**.*

*Other factors may have contributed to **measurement error**.*

The second wrong:

MEASUREMENT ERROR

Polls Are Not Machines

Catch-all for Mistakes and Assumptions

PROBLEMS INCLUDE:

- INCORRECT SAMPLE SELECTION
- RESPONDENT MISTAKES
- DATA TABULATION OR PROCESSING MISSES



Mainstream Media Accountability Survey

1. Do you believe that the mainstream media has reported unfairly on our movement?

- ☐ Yes
- ☐ No
- ☐ No opinion

2. Do you trust MSNBC to report fairly on Trump's presidency?

- ☐ Yes
- ☐ No
- ☐ No opinion

3. Do you trust CNN to report fairly on Trump's presidency?

- ☐ Yes
- ☐ No
- ☐ No opinion

17. Do you believe that the media has been far too quick to spread false stories about our movement?

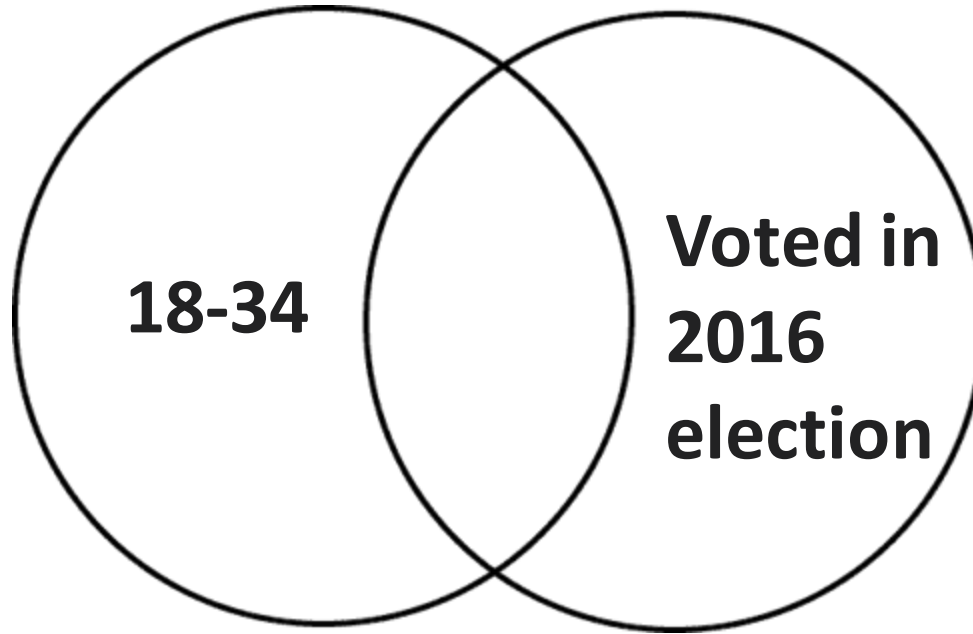
- ☐ Yes
- ☐ No
- ☐ No opinion
- ☐ Other, please specify:

18. Do you believe that the media uses slurs rather than facts to attack conservative stances on issues like border control, religious liberties, and ObamaCare?

- ☐ Yes
- ☐ No
- ☐ No opinion
- ☐ Other, please specify:

Coverage Error (Sample Selection)

- When the survey talks to the wrong population.



Coverage Error (Sample Selection)

HOW TO SPOT:

- Look at the specified sample and ask, is this the right group?
- Does this population make sense with the questions?
- Can a survey actually reach the populations they claim to be covering.

Response Errors

- When respondents provide incorrect information.
- Intentionally: Respondents misrepresent themselves.
- Unintentionally: Respondents make mistakes answering questions.

**HUFFPOLLSTER: Battleground State Polls Missed
In White Working-Class Areas**

SHY TRUMPERS

**The rich and educated don't like to admit
they're voting for Trump**

How Many People Support Trump but Don't Want to Admit It?

Response Errors

HOW TO SPOT:

- **Intentional** is hard to spot.
 - Analysts have to check for discrepancies in answering.
 - Reporting and investigation is actually a great way to detect.
- **Unintentional** is about being able to understand the perspective of respondents.
 - Are the questions understandable and clearly designed?
 - Is the topic something people know enough to answer?

Data Tabulation Errors

Mistakes when tabulating data

- Less of an issue with technology adoption

HOW TO SPOT:

- For polling in developed economies, less of an issue.
- For polling face to face or in developed markets, ask about quality control standards

Poorly Worded Questions

Related to respondent error

Researchers may, intentionally or unintentionally, write bad questions

EXAMPLES INCLUDE:

- Double barreled questions
- Overly technical questions
- Leading questions
- Incorrect response options

Q. Which do you think is better for the country? Should the Democrats and Republicans compromise some of their positions in order to get things done, or stick to their positions even if it means not getting as much done?

Compromise
Stick to positions
Don't know

Source: New York Times/CBS News Poll 1/12-17/2012

The third wrong **INTERPRETATION ERROR**

Even with a trustworthy poll, errors can come in how we as journalists interpret the data

Seven ways to avoid interpretation errors

1. You can't always invert the pie
2. Pay attention to the sample (who's in, and who's out?)
3. Data you see might not come from a trained data viz expert
4. A small change might not be a change at all, or might be opposite what you think it is
5. Beware of what Tufte calls "Spurious precision"
6. Generalize: "most," "nearly," "roughly," and "about" are your friends
7. Be especially careful of sub-groups

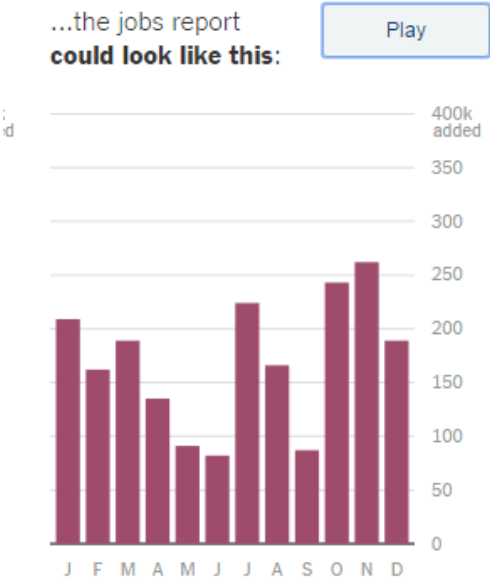
Welcome to Bullock Co.: Hispanic pop. -9



« < 1 - 18 of 6,284 > »

	Bullock County, Alabama	
	Estimate	Margin of Error
Total:	10,678	*****
Not Hispanic or Latino	10,205	+/-482
Hispanic or Latino	473	+/-482

Watch the margins

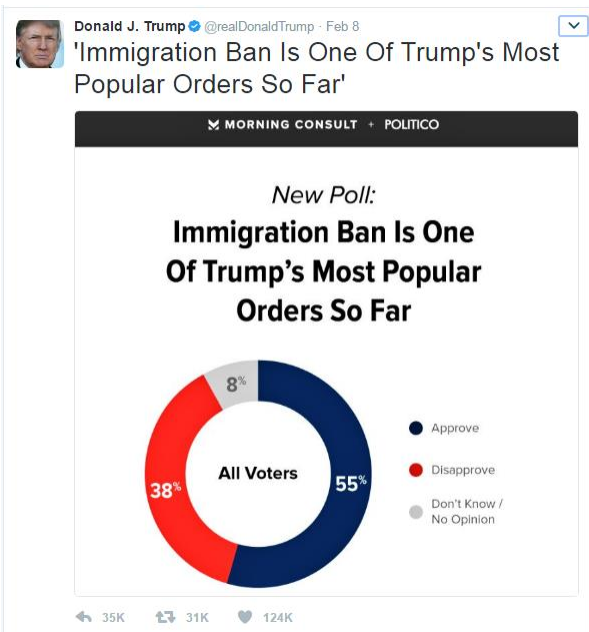


If the economy actually added 150,000 jobs last month, it would be possible to see any of these headlines:

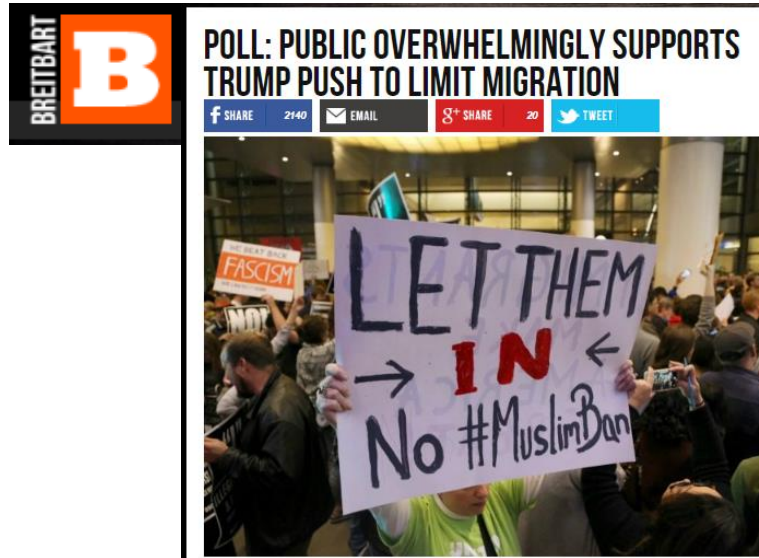
The jobs number is just an estimate, and it comes with uncertainty.



Avoid cherry-picking



Don't bend a poll to your will



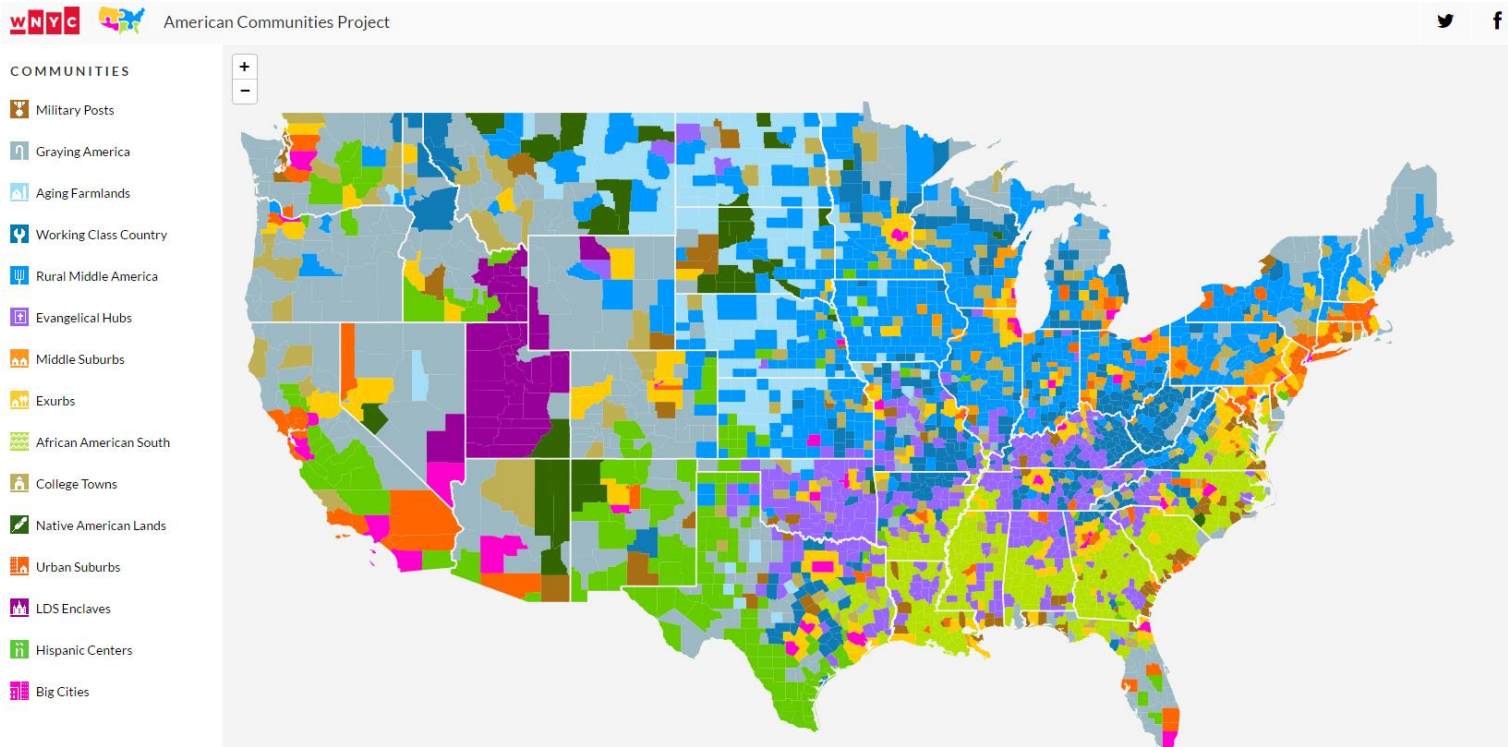
When in doubt: Report

And now, some rights:

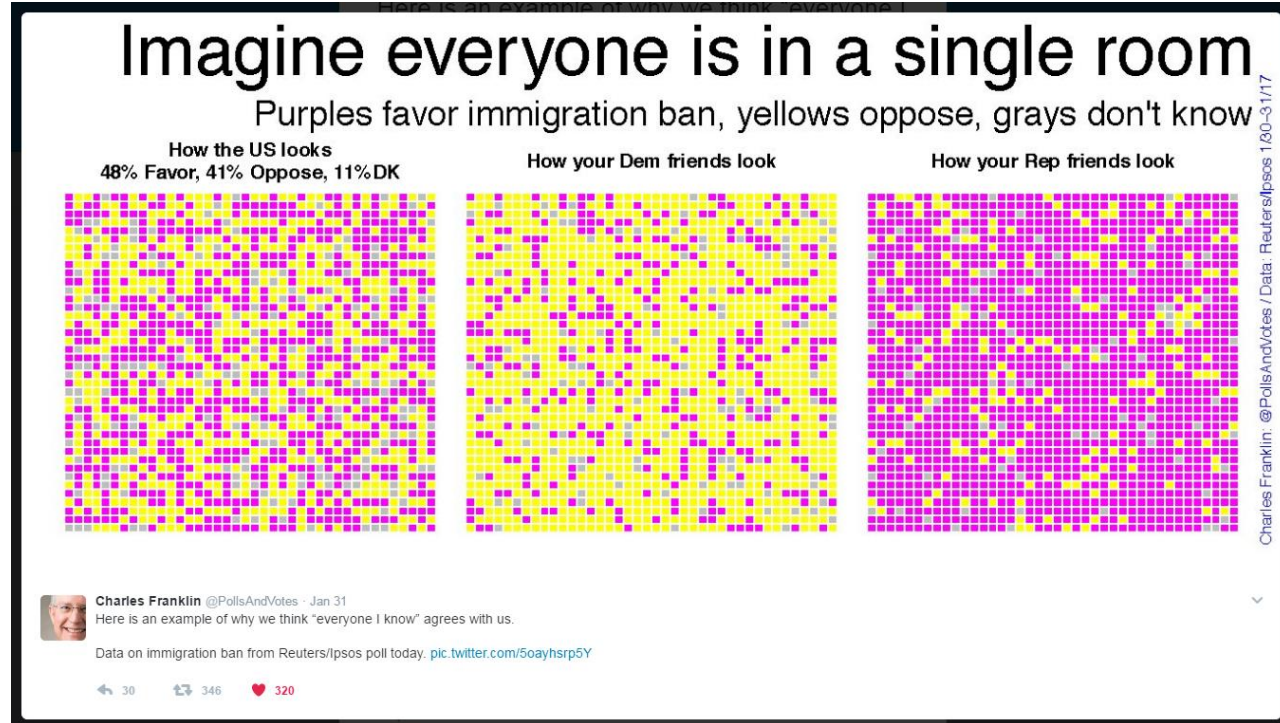
WAYS TO USE SURVEYS

Some quick examples of creative ways to use polling and survey data to add context and texture to your reporting

Choose your own adventure

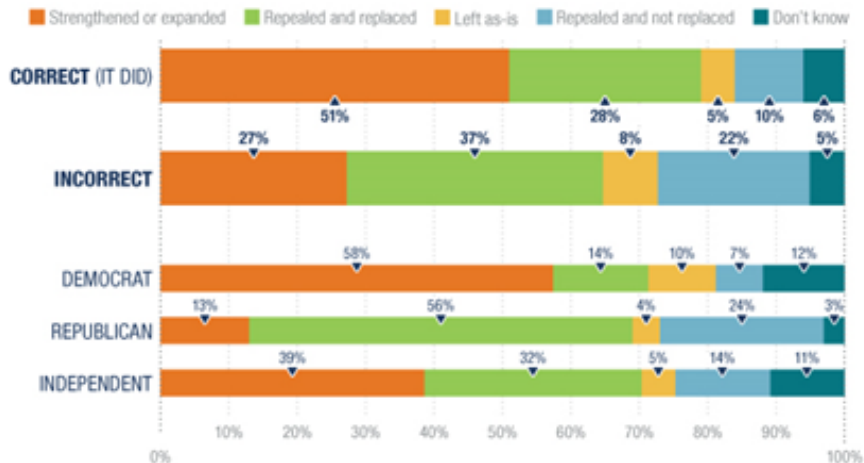


Make it your own

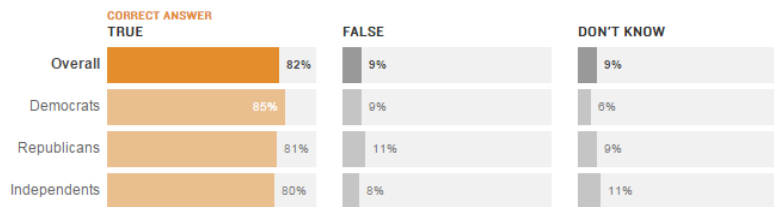


Take a different approach

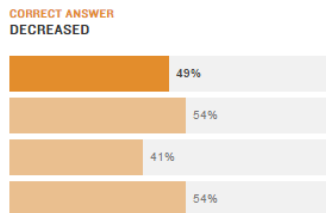
Q: The ACA lead to a decrease in the number of uninsured Americans



True or false: On average, Americans pay more for health care than people in other countries



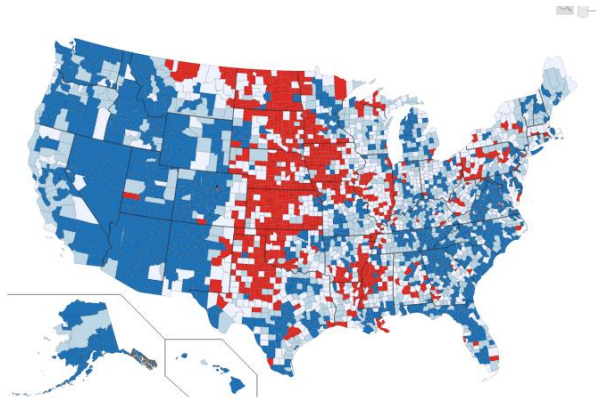
Since the passage of the ACA, has the number of uninsured Americans increased, decreased or stayed the same?



Bake it into something else

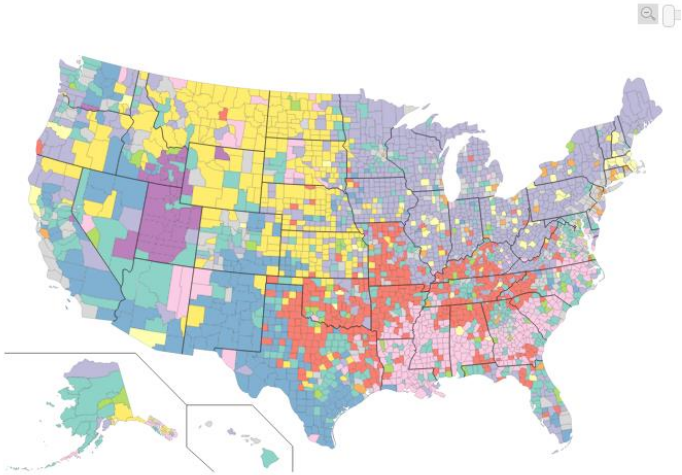


Asking nicely works



Source: U.S. Census Bureau

Top Network Shows by County Type



Source: Experian Simmons

- Raising Hope (FOX)
- 30 Rock (NBC)
- Mike & Molly (CBS)
- Supernatural (The CW)
- Frontline (PBS)
- America's Next Top Model (The CW)
- 90210 (The CW)
- Cold Case (CBS)
- Modern Family (ABC)
- Chuck (NBC)
- Blue Bloods (CBS)

Thank you!



Matt Carmichael

Dir. Of Editorial Strategy
Ipsos NA

✉ Matt.carmichael@ipsos.com



Peter Gade

Gaylord Chair of Journalism
University of Oklahoma

✉ pgade@ou.edu



Chris Jackson

Vice President
Ipsos Public Affairs

✉ Chris.Jackson@Ipsos.com