

Deciphering the Signal and the Noise

Understanding How FiveThirtyEight's Election Forecast Works
and Why It Might Have Failed in the 2016 Presidential Election

By Benjamin Tanen

EM-52: Technical and Managerial Communication
Tufts University

Overview

1. Introduction
2. How the Election Forecast Works
3. Audience Considerations for Election Forecasts

Purpose

Demystify the technical aspects of FiveThirtyEight's 2016 Election Forecast so the audience might feel more comfortable with predictive election models in general

Background on FiveThirtyEight and their Election Forecast

FiveThirtyEight is an online organization for data-driven analysis on political and sports news.

Nate Silver formed the site after he successfully predicted the 2008 election.

After successful models in 2008 and 2012, FiveThirtyEight published an updated Election Forecast for the 2016 presidential election.



 **FiveThirtyEight**

Historical Performance of FiveThirtyEight's Forecast

In 2008, Silver correctly predicted results across the U.S. except for Indiana and one district in Nebraska.

In 2012, FiveThirtyEight correctly predicted all 50 states and the District of Columbia.

In 2016, FiveThirtyEight predicted Hillary Clinton would win by a decent margin. This was incorrect.

How the Election Forecast Works

1. Data Sources the Forecast Relies On
2. Weighted Average Calculations
3. Individual Poll Adjustments
4. Additional Adjustments
5. Simulating the Election

Data FiveThirtyEight Uses in Its Election Forecast

- Outside polling data
 - Excludes any polls conducted by a campaign, PAC, Super PAC, or banned pollster
- FiveThirtyEight's Pollster Ratings
- Economic indicators
 - Metrics for (1) strength of the jobs market, (2) strength of manufacturing, (3) average U.S. income, (4) consumer spending, (5) inflation, and (6) stock market performance

Calculation of Weighted Averages

Using all of the included polls, the model calculates weighted averages for each of the 50 states plus the District of Columbia.

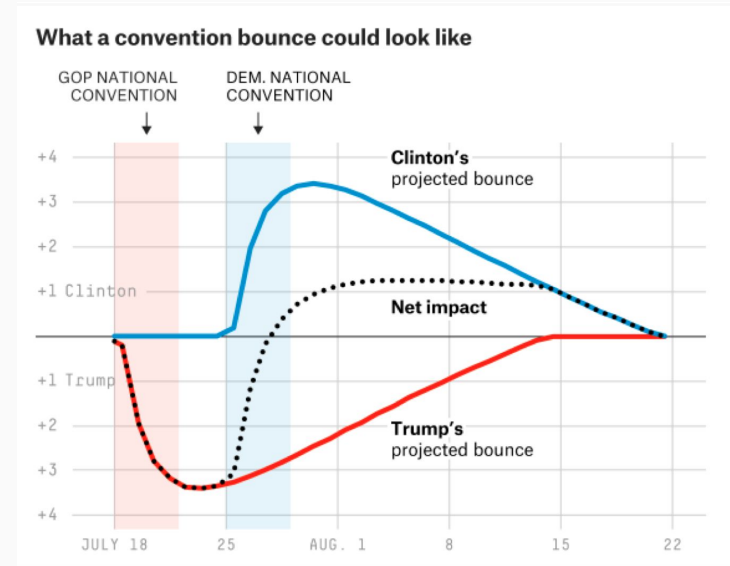
Each poll's weight is determined by:

- The pollster's rating
- The sample size
- The recentness of the poll

Individual Poll Adjustments

After aggregating the polls, FiveThirtyEight makes a number of adjustments to the averages. These include:

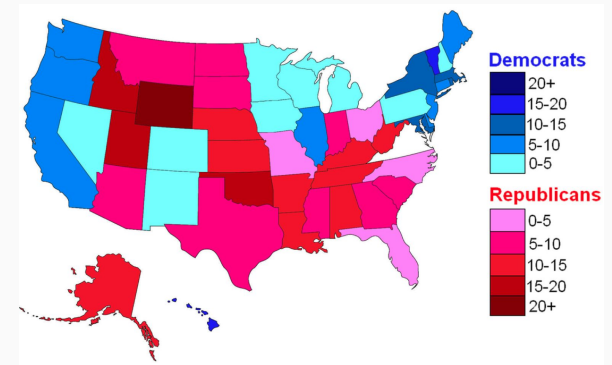
- A convention bounce adjustment
- A “missing Johnson” adjustment
- A pollster bias adjustment



Adjusting for Polling Demographics

To account for demographic biases, FiveThirtyEight makes further overall adjustments using Partisan Voter Indices (PVI).

PVIs are measures of how much a demographic group might lean either Democratic or Republican.



State PVI from 2012 election

Adjusting for Economic Conditions

FiveThirtyEight also added a measure for the strength of the U.S. economy.

The metric was calculated as a normalized average of the six previously mentioned economic indicators.

If the economy was *relatively* strong, Clinton would receive a small boost in her predicted performance.

Simulating the Election

After aggregating and adjusting the polls for each state, the 51 averages (one for each state plus D.C.) were combined into one score for each candidate.

These averages were updated ~20,000 times a day on FiveThirtyEight's site.



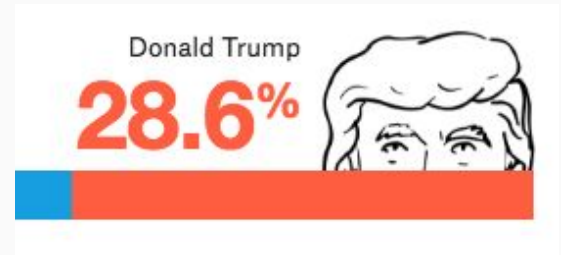
Audience Considerations for Election Forecasts

1. The Probabilistic Nature of Election Forecasts
2. The Inherent Error Present in Election Polling
3. The Uncertainty in Modeling

The Probabilistic Nature of Election Forecasts

FiveThirtyEight projected Trump had a 28.6% chance of winning, yet many people took that to mean an absolute impossibility.

Most models are based on probabilities, so it is important to recognize their predictions are not guarantees.



The Inherent Error in Election Polling

Polls use statistical sampling, which **always** include some margin of error. This means when polls are aggregated, their errors also aggregate.

If every poll is wrong “in the same direction,” forecasting models will also fail in that same direction.

Readers should always remember outside contributors to error.

The Uncertainty of Modeling (and Being Okay with It)

Forecasting models cannot account for everything, some things just remain intangible.

“The world is largely unknowable, so we just try to take what we can know and make the most of it” - Nate Silver, 2017

Conclusions

If a model fails, it is not necessarily because it is a bad model - good models fail too!

To distinguish between the good and the bad, readers should consider what a model predict **and** how it makes these predictions.

Predictive modeling is not a fool's errand.