

PredictWise

Overview

Major advantages of our methodological approach:

1. **Accurate:** Our state-of-the-art method uses machine learning to convert whatever polling data we collect into results that are at least as accurate as results from traditional random-digit dialing polls. We have showcased accuracy of these methods extensively in a validation study with PEW ([here](#)), and in this academic study ([here](#)).
2. **Depth:** We can present movements with unprecedented demographic and geographic granularity. We usually project for more than 150 demographic categories, but in theory we can project down to the Census Block group.
3. **Speed:** We can assess opinions on political matters within approximately three hours from customer request to result presentation. This also allows us to assess public opinion to crises relating to core democratic institutions, protests, etc.
4. **Cost:** This method costs us about 5 to 10% of traditional RDD polling and raking.

Methodology

Our method centers on state-of-the-art deep and machine learning techniques, leveraging the rich array of passive data exclusive to mobile data collection. While the raw survey data we obtain from mobile polls are of high quality, we refine representativeness using computational algorithms we have carefully fine-tuned over the last two years.

First, we model the raw responses to each question, given each respondent's age, gender, location, education level, race, marital status, party identification, income, family size and urbanicity derived from the geo-identified residential location. This information divides the population into thousands of demographic categories. For each sub-group and poll question, we predict the percent of people that would provide each answer if the entire country showed up to the poll.

Each of these predictions is informed by all responses, including responses received in previous iterations of the poll. To achieve this, we have developed a complex dynamic model that allows us to parse out variance in sample composition from true swings over time. This is crucial for assessing trends in polling data, as illustrated by [this cautionary tale](#) from the world of political polling.

Second, we project our estimates for each sub-group onto our best estimate of the likely adult population, a process known as post-stratification. Specifically, we weight our predictions by the fraction of that sub-group in the overall target population. We derive the target population from a Big Data combination of population-level census data, and proprietary financial and political data sets providing information on every registered voter in the US.

The transformed data provides meaningful information about many segments of the population. Our approach has been validated in peer-reviewed articles published in leading academic journals (see [here](#) and [here](#)) and real-world event predictions (see [here](#) and [here](#)). Our polling has lead to major publications in the US' leading news platforms, including but not exclusive to *The New York Times* (see [here](#) and [here](#)), the *Washington Post* (see [here](#) and [here](#)), or *Slate* (see [here](#)).