

# Opinion Poll - Influencing factors, methodologies and their accuracies

## 1 Introduction

A healthy public polling system is a sign of a free society. It demonstrates the ability of non-government organizations to collect and publicize information regarding the public's well-being and citizens' perspectives on significant topics. In countries where polling is scarce, the head of state can simply proclaim residents' wishes and requirements.

There are various factors that influence opinion polls. This paper takes a deep dive into some of those critical elements and how they influence the surveys. The paper also discusses the various methodologies employed by polling organizations in conducting opinion polls and their impact on the results. A small piece of analysis has been done to demonstrate the same.

## 2 Background

Election polling is perhaps one of the most common types of statistics in everyday life, and it is gaining in popularity in both the popular press and public debate. While polling dates back several millennia to censuses and straw polls done early in America's political history, it was not until the twentieth century that surveys became more commonly used (Weisberg 2009). Since the 1990s, the number of public polls has continued to rise, with polling becoming firmly embedded in campaign news stories by the early 2000s (Frankovic 2005).

The most pivotal event in opinion polling history came in 1936, when President Franklin D. Roosevelt faced Republican candidate Alf Langdon in the presidential election. The Literary Digest conducted a survey by mailing out around 10 million postcards to persons registered in the phone book or on automobile owner registries, as had been done previously and successfully. The magazine got 2.3 million responses after asking readers to vote for their preferred presidential candidate. The magazine predicted a Landon win of 57% to 43% of the two-candidate popular vote, but Roosevelt won by a landslide, garnering 62.5 percent of the vote in the election. Following this, the Literary Digest lost its reputation and became insolvent, resulting in a 1938 merger with Time magazine (Macreadie 2011).

In terms of polling, George Gallup came out on top in 1936, correctly predicting Roosevelt's huge victory. Gallup's own door-knock survey showed Roosevelt

winning 56-44, which underestimated the President's victory margin but was considerably closer to the actual election outcome. Gallup's method of random sampling to assure accuracy was a key to his success in 1936. His success challenged the contemporary methodologies in opinion polling and had a considerable impact on the development of survey research techniques not just in the United States, but across the world.

Quantitative opinion polling or qualitative survey research can be used to gauge public opinion on political topics.

## 2.1 Quantitative survey

Surveys that quantify the opinions of a group of individuals are referred to as quantitative opinion polling. These are especially handy in election circumstances with straightforward questions and answers limited to yes/no. Questionnaires, face-to-face interviews, telephone surveys, and online/email surveys are often used in quantitative opinion polling.

## 2.2 Qualitative survey

Focus groups, in-depth interviews, and participant observation are common components of qualitative survey research. Internal opinion research by political parties and lobbying organizations may help them better understand the causes for public mood, as well as how individuals perceive politicians, programmes, and parties through qualitative research. A qualitative survey may elicit more personal information from respondents, which may help political parties focus campaigns and appeal to certain voter groups. The most popular type of qualitative survey research is focus groups, which are made up of small groups of individuals who participate in a controlled and recorded conversation.

## 3 Influencing factors

There are several factors and conditions that need to be taken into consideration while conducting opinion polls. The sorts of questions asked and the sequence in which they are asked, the sample size, the methodology used, the timing of polls in connection to elections or events, and many other elements and circumstances, restrictions, and reservations all impact the outcome of opinion surveys.

### 3.1 Margin of error

The margin of error (also known as sampling error) is a statistic that expresses the degree of random sampling error in the results of an opinion poll. Pollsters often claim the margin of errors in their conducted polls as  $\pm 3\%$  which lead people to think that polls are more precise than they really are. A 3% point margin of error indicates that a poll result for any candidate may be 3 points wrong in either way. A 1,000-person survey has a margin of error of  $\pm 3\%$ , while a 2,000-person poll has a margin of error of  $\pm 2\%$ . The smaller the sample, the less exact it is, and the greater the error margin. These calculations assume that polls are true random sampling, with every member of the population having an equal probability of being picked. However, this is rarely true. Polls are conducted using quota sampling or from panels of volunteers. Sample sizes are significantly smaller for smaller demographic groupings, resulting in substantially higher margins of error. A survey of 1000 individuals, for example, would have a margin of error of  $\pm 3\%$ , but if only 100 Hispanics responded, the Hispanic statistics would have a margin of error of  $\pm 10\%$ . This means it would be only be statistically significant if the difference between the responses from the Hispanics and that from the rest of the sample was more than 10%. The error is particularly common when taking into account the responses from ethnic minorities or religious minorities in national polls. However, when compared to genuine events like general elections, the margin of error is still a fair approximate indicator to how exact a poll is, and most polls are within the margin of error of the true result.

### 3.2 Sample size and representative samples

The square root of the sample size determines the margin of error (Gelman 2004). A sample size of 250 will give a margin of error of 6%, whereas a sample size of 100 will give a margin of error of 10%. A 10% margin of error is not efficient. On the other hand, by surveying 4000 people the margin of error can be brought down to 1.5 percent. However, since public opinion varies from day to day, it is meaningless to attempt such precise estimate. Hence a reasonably sized sample is ideal. The margin of error is a mathematical construct, and actual survey inaccuracies are bigger for a variety of reasons. Even when random sampling is used, people in the population have uneven chances of being selected for the survey. For instance, people without telephones will not be included in the survey and people who have more than one telephone will have more chances of being included in the survey. If substantial numbers of a demographic category, such as young people or individuals who do not speak English, are not included in a sample, the sample is not representative. It's challenging to get representative samples. Due to declining response rates, polling certain demographics, such as those who favor certain political parties, has become more challenging. These biases are corrected

by modifying the sample to match the population, but such adjustments can never be perfect because they only correct for recognized biases.

### 3.3 Addressing Uncommitted responses

When it comes to dealing with "don't know," "uncommitted," and "no" replies, different polling firms have varied methodologies. Some firms may choose to ignore these replies in their computations, while others may choose to distribute responses based on the respondents' political leanings. Some polling organizations will distribute "don't know" replies in accordance to the reported estimate of support for each party, rather than political leaning.

Gallup discovered many flaws in public opinion polling, such as:

- It is possible that the respondents have no knowledge of the topic under discussion.
- There seems to be no distinction made between those who make snap decisions and those who have carefully considered the advantages and drawbacks of a situation.
- Although most responses are classified as 'Yes' or 'No,' certain complicated topics cannot be reduced to a single, dichotomous question (Gallup 1947).

Excluding uncommitted and non-responses from public polls to estimate election outcomes may make particular sense in countries where voting is not mandatory, because those who don't have an opinion or have no interest in politics are less likely to vote.

Gallup suggested that people with no prior knowledge of the subject be removed from surveys. Gallup devised a "quintamensional" method to question design to facilitate this, which was based on five types of questions in measuring opinion, the first of which was a "filter" question that aimed to remove individuals who had no prior knowledge of the subject. While the other categories were used to determine the intensity of opinion and the reasons for it.

### 3.4 Opinion Strength

Quantitative opinion surveys frequently fail to account for the strength of a person's feelings about a topic or candidate. Opinion polls group responses into categories, and respondents are frequently limited to one of two or more options. Opinion surveys convert those being examined to objects with bundles of qualities in

order to create an instrumental result by attempting to categorize them and their opinions(Dryzek 1990).

In order to avoid uncommitted replies, polling organizations may ask respondents to choose which party they have a higher "leaning toward" when it comes to political party questions. Respondents have the option of selecting an alternative, however their replies are weighted equally with those who have strong feelings about a political party.

In areas where voting is not mandatory, polling organizations confront extra hurdles. Pollsters may ask respondents if they intend to vote before determining their view, but the respondent may be intending to vote but still uncertain, or the respondent may be extremely susceptible to political rhetoric in the closing days of a campaign(Zurkin 2004).

### 3.5 Weighting

The poll data is weighted in order to ensure that the sample more precisely represents the characteristics of the population from which it was collected and from which an inference will be formed. Weighting is a method of adjusting the relative contribution of respondents without affecting the actual responses to survey questions.

Weights can be used to -

- Adjust for the likelihood of a respondent being chosen in a survey.
- Make adjustments to account for the sample design's characteristics.
- After the data has been gathered, make modifications to put certain characteristics of the sample in line with other known population characteristics.

#### 3.5.1 Adjusting for probability

Every participant must have a known, non-zero chance of being chosen, according to one of the basic principles of probability sampling. The sample in a typical media survey is meant to reflect all individuals aged 18 and up. Because of the way the data is obtained, this also implies "among those dwelling in telephone homes." As a result, survey data is weighted to account for the likelihood of picking a respondent.

### 3.5.2 Adjusting for sample design

To enhance the sample size, stratified sampling is a type of design in which groups of prospective responders with an identifiable feature are picked with uneven weights. An oversample is a term used to describe this situation. It's utilized to get more exact estimations of how that group feels about the campaign or how they're reacting to it.

For example, African Americans, who make up approximately 12% of the US population, were oversampled in certain surveys during the 2008 presidential campaign to gauge their attitudes to Barack Obama and Hillary Clinton. The reactions of whites and African Americans may be easily compared using data. However, if a national estimate of Clinton and Obama support is required, a weight must be devised to “**reduce**” the contribution of each African American respondent while “**increasing**” the contribution of each white respondent.

### 3.5.3 Adjusting for demographics

Finally, once all of the data has been gathered, certain simple frequencies are applied to specific demographic factors so that they may be compared to known population characteristics received from an external source, such as the United States Census. Because some demographic groups are over-represented or underrepresented in the sample, this adjustment is applied.

Young males, for example, are far more difficult to reach at home than elderly ladies. As a result, unweighted data often has a higher proportion of older women and a lower proportion of younger males than the U.S. Census reports. Small adjustments, known as post-stratification weights, are often made by pollsters to bring the sample into line with known demographic characteristics such as age, gender, location, and education.

## 3.6 Question Design

Question design is a critical element that can lead to discrepancies between polling organizations' results on the one hand, and between opinion polls and actual election outcomes on the other. There had been far too much focus on diverse sampling procedures to explain differences in findings, and far too little on question design (Gallup 1947).

Many factors influence question design, including the wording and tone of the questions, the quantity of questions asked, the sequence in which they are asked,

the usage of important words, such as emotional, biased, or suggestive language, and the values embedded in the questions asked, to mention a few. Opinion polls frequently pose particular questions on a specific topic, which may or may not be important to the interviewee, but which, by their very nature, have the potential to affect respondent attitudes about the relevance and worth of such issues.

## 4 Methodologies used in opinion polls

Before adapting to the growth of telephone connection, the polling business relied on mail and face-to-face interviews. It is undergoing yet another change, this time to accommodate the widespread availability of internet connection. This indicates we're in the midst of an era marked by a wide range of survey methodologies such as the below –

- **Live telephone:** A real person conducting a telephone conversation with the respondent to gather their replies to the questions asked.
- **Interactive voice response (IVR):** Respondents engage with a recorded voice and type or speak their replies into a keypad.
- **Internet:** Respondents answer survey questions utilizing the internet on a computer, tablet, or mobile.
- **IVR + Internet:** A mix of both the IVR and Internet methods mentioned above.
- **IVR + Live telephone:** A combination of the IVR and live phone techniques mentioned above. Because of the greater response rate, live interviewers are regarded the gold standard in survey research.

In this section and the subsequent ones, we will discuss the changes in popular methodologies over a period of time as well as assess their accuracies using the raw dataset obtained from the database of the polling aggregator company FiveThirtyEight which contain polling data accumulated from different polling organization related to polling held the United States for elections.

The raw dataset comprises of 27 attributes (Silver 2014). However, for our piece of analysis we would only need the columns mentioned with their descriptions in table 1.0 below. It may also be worth mentioning that the data in this dataset is at the most granular level with a poll-id for each poll conducted. The data satisfy the criteria –

1. The polls were conducted in the year 1998 or later.

2. The polls were conducted in the last three weeks of campaign.
3. The polls were conducted in one of the following types of elections.
  - Presidential general elections
  - Presidential primaries
  - Senate elections
  - Gubernatorial elections

Feature Name	Definition
year	Year of election (not year of poll)
location	Location (state or Congressional district, or "US" for national polls)
pollster	Pollster name
methodology	Live Phone, IVR, Mail, Online, Text, Face to Face. <small>*Mixed method polls list the methods involved in the poll separated by slashes.</small>
cand1_name	In races where a Democrat and a Republican were the top two finishers, Candidate #1 is the Democrat.
cand1_pct	Candidate #1's share of the vote in the poll.
cand2_name	In races where a Democrat and a Republican were the top two finishers, Candidate #2 is the Republican.
cand2_pct	Candidate #2's share of the vote in the poll.
margin_poll	Projected margin of victory (defeat) for Candidate #1.
cand1_actual	Actual share of vote for Candidate #1
cand2_actual	Actual share of vote for Candidate #2
margin_actual	Actual margin in the election. This is calculated as $\text{cand1\_actual} - \text{cand2\_actual}$ .
error	Absolute value of the difference between the actual and polled result.
bias	Statistical bias of the poll. This is calculated only for races in which the top two finishers were a Democrat and a Republican. It is calculated as $\text{margin\_poll} - \text{margin\_actual}$ . Positive values indicate a Democratic bias (the Democrat did better in the poll than the election). Negative values indicate a Republican bias.
rightcall	Flag to indicate whether the pollster called the outcome correctly

Figure 1: Table 1.0

The data in the dataset was aggregated by methodologies and year in order to observe the changes in the popularity of these methodologies over the years. The mixed category comprises of polls conducted using mixed methods which are - 'IVR/Live Phone', 'IVR/Online', 'Live Phone/Online', 'IVR/Live Phone/Text', 'IVR/Live Phone/Online', 'IVR/Text', 'Text', 'IVR/Online/Text', 'IVR/Live Phone/Online/Text', 'Live Phone/Text', 'Online/Text', 'Live Phone/Online/Text'



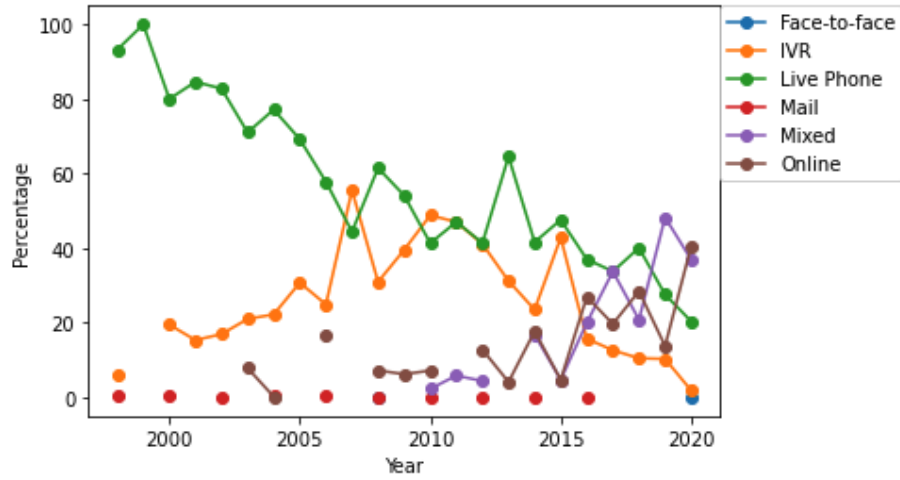


Figure 2: Changes in polling methodologies over years.

It can be observed in the above figure that the popularity of polls conducted via live phones was much higher in the initial part of this time period and the popularity reduces as approach the current years. The popularity of IVR can also be observed to have reduced in the recent years. Whereas the popularity of polls conducted online and via the mixed category has been observed to grow in the recent years.

#### 4.1 Error in methods

In order to understand which polling methods perform better we analyzed the data using the error count for each poll. Error is calculated as the absolute value of difference between the actual result and the polled result. The following boxplot of errors by methodologies has been plotted using the same.

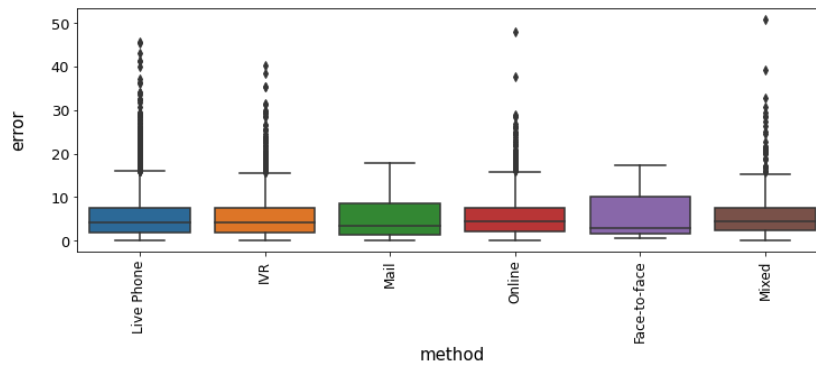


Figure 3: Boxplot of error by methodologies.

It can be observed from the box plots above that the median error in all the

methodologies are nearly the same and ranges around 4 to 6 points. A long tail with outliers can be seen in the box plots of methodologies – Live Phone, IVR, Online and Mixed. This could be attributed to the fact that more number of polls are conducted through these methodologies as compared to the rest and hence has a higher chance of having potential outlier.

## 4.2 Correct predictions by methodologies

A poll is called in right if the polling organization is able to rightly predict the outcome of the actual election. The ‘right call’ attribute in the dataset is an indicator column which holds the value 1 if the pollster is able to predict the outcome correctly and 0 otherwise. The below bar graph tells us about the percentages of polls predicted correctly by the different methodologies.

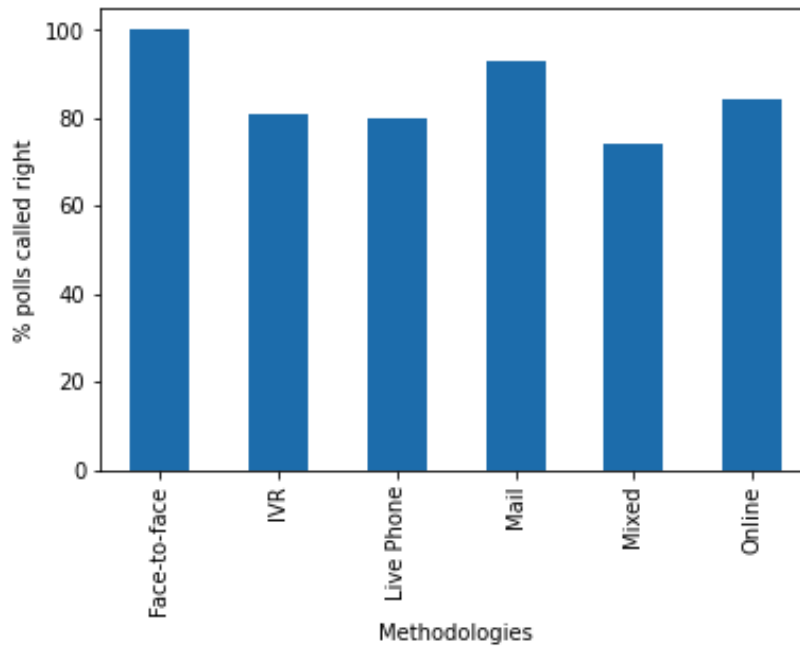


Figure 4: Percentage of correctly predicted polls by methodologies.

## 5 Bias

The raw dataset used for this analysis allows us to identify potential bias in polls. When compared to actual election outcomes, the word bias refers to how much a polling firm’s results leaned toward one party or the other. In the dataset, this is calculated only for those races in which the top two finishers were a Democrat or a Republican. Bias is calculated as the difference in the poll result and the actual

result. In our dataset, a positive value for bias indicates a democrat bias whereas a negative value indicates a republican bias. In other words, if a pollster had the democrat candidate ahead in the poll by 5 points whereas the actual election result was a tie, the pollster has a bias towards Democrats but if the Democrats win the election by 10 points, the pollster has biased towards Republicans by 5 points.

In order to see if the pollsters with certain partisan bias are more successful in predicting the election outcomes correctly, we generated the below box plots to the data aggregated at the location level. The aggregation at location level has been used with the intuition that a pollster's bias towards a particular party may vary from state to state.

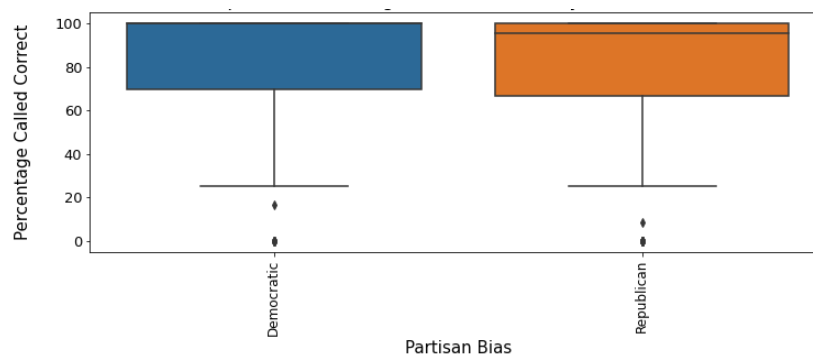


Figure 5: Boxplot of Percentage Called Correct by Partisan Bias

It can be observed from the plot above that the Democratic leaning pollster have been able to call the elections more correctly than the Republican leaning pollsters in their respective states where they favor one party over the other.

## 6 Conclusion

If phone is losing popularity, what is filling its void? Our research from above tells us it is the online internet polls.

As more people gained access to the internet, pollsters began to search for a means to reach respondents via the internet. There are several advantages to using this approach. People may do the survey at their convenience, pollsters don't have to employ and manage a roomful of live interviewers or pay phone expenses, and survey methodologists have discovered that self-administration has measurable advantages.

However, there is one big major obstacle. While it is possible to generate random samples of the US population using master lists of people's home

addresses or phone numbers (through the U.S. Postal services and the Federal Communications Commission, respectively), there is currently no means to do it over the internet(Dean 2019).

This is significant because traditional survey research relies heavily on the statistical theory of the random sample, according to which every member of the population has an equal (or at least known and nonzero) probability of being included. This results in polls that accurately represent the country's racial, ethnic, religious, and socioeconomic diversity. Low response rates can skew the sample's unpredictability, but there's evidence that starting the polling process with a genuine cross-section of the population improves accuracy.

Our analysis on pollster's partisan bias could raise a question in the reader's mind regarding the neutrality of polling organizations. Different experts in this field have different opinions regarding this and a common solution seems to be survey aggregators. Survey aggregators assist in bringing together a wide range of data. There is definitely something helpful about looking at a range of data which help data scientists to draw trends, look for patterns and valuable insights.

## References

- Dean, Claudia (2019). *A Field Guide to Polling: Election 2020 Edition*. URL: <https://www.pewresearch.org/methods/2019/11/19/a-field-guide-to-polling-election-2020-edition/>.
- Dryzek, J. (1990). "Discursive Democracy". In: *Political Campaign Strategy: Doing Democracy in the 21st Century*.
- Frankovic, K.A. (2005). "Reporting "The Polls"". In: *Public Opinion Quarterly* 5, pp. 682–697. DOI: 10.1093/poq/nfi066.
- Gallup, G. (1947). "The Quintamensional Plan of Question Design". In: *Public Opinion Quarterly* 11, pp. 385–393.
- Gelman, Andrew (2004). *How can a poll of only 1,004 Americans represent 260 million people with only a 3 percent margin of error?* URL: <https://www.scientificamerican.com/article/howcan-a-poll-of-only-100/>.
- Macreadie, Rachel (July 2011). "Public Opinion Polls". In: DOI: 10.13140/2.1.2546.4646.
- Silver, Nate (2014). *How FiveThirtyEight Calculates Pollster Ratings*. URL: <https://fivethirtyeight.com/features/how-fivethirtyeight-calculates-pollster-ratings/>.
- Weisberg, H.F. (2009). *The total survey error approach a guide to the new science of survey research*. University of Chicago Press.
- Zurkin, C. (2004). "Sources of Variation in Public Election Polling: A Primer". In: *American Association for Public Opinion Research*.