

Case Study I: The 1936 *Literary Digest* Poll

The presidential election of 1936 pitted Alfred Landon, the Republican governor of Kansas, against the incumbent President, Franklin D. Roosevelt. The year 1936 marked the end of the Great Depression, and economic issues such as unemployment and government spending were the dominant themes of the campaign. The *Literary Digest* was one of the most respected magazines of the time and had a history of accurately predicting the winners of presidential elections that dated back to 1916. For the 1936 election, the *Literary Digest* prediction was that Landon would get 57% of the vote against Roosevelt's 43% (these are the *statistics* that the poll measured). The actual results of the election were 62% for Roosevelt against 38% for Landon (these were the *parameters* the poll was trying to measure). The sampling error in the *Literary Digest* poll was a whopping 19%, the largest ever in a major public opinion poll. Practically all of the sampling error was the result of sample bias.

The irony of the situation was that the *Literary Digest* poll was also one of the largest and most expensive polls ever conducted, with a sample size of around 2.4 million people! At the same time the *Literary Digest* was making its fateful mistake, George Gallup was able to predict a victory for Roosevelt using a much smaller sample of about 50,000 people.

This illustrates the fact that bad sampling methods cannot be cured by increasing the size of the sample, which in fact just compounds the mistakes. The critical issue in sampling is not sample size but how best to reduce sample bias. There are many different ways that bias can creep into the sample selection process. Two of the most common occurred in the case of the *Literary Digest* poll.

The *Literary Digest's* method for choosing its sample was as follows: Based on every telephone directory in the United States, lists of magazine subscribers, rosters of clubs and associations, and other sources, a mailing list of about 10 million names was created. Every name on this list was mailed a mock ballot and asked to return the marked ballot to the magazine.

One cannot help but be impressed by the sheer ambition of such a project. Nor is it surprising that the magazine's optimism and confidence were in direct proportion to the magnitude of its effort. In its August 22, 1936 issue, the *Literary Digest* announced:

Once again, [we are] asking more than ten million voters -- one out of four, representing every county in the United States -- to settle November's election in October.

Next week, the first answers from these ten million will begin the incoming tide of marked ballots, to be triple-checked, verified, five-times cross-classified and totaled. When the last figure has been totted and checked, if past experience is a criterion, the country will know to within a fraction of 1 percent the actual popular vote of forty million [voters].

There were two basic causes of the *Literary Digest's* downfall: selection bias and nonresponse bias.

The first major problem with the poll was in the selection process for the names on the mailing list, which were taken from telephone directories, club membership lists, lists of magazine subscribers, etc. Such a list is guaranteed to be slanted toward middle- and upper-class voters, and by default to exclude lower-income voters. One must remember that in 1936, telephones were much more of a luxury than they are today. Furthermore, at a time when there were still 9 million people unemployed, the names of a significant segment of the population would not show up on lists of club memberships and magazine subscribers. At least with regard to economic status, the *Literary Digest* mailing list was far from being a representative cross-section of the population. This is always a critical problem because voters are generally known to vote their pocketbooks, and it was magnified in the 1936 election when economic issues were preeminent in the minds of the voters. This sort of sample bias is called *selection bias*.



The second problem with the *Literary Digest* poll was that out of the 10 million people whose names were on the original mailing list, only about 2.4 million responded to the survey. Thus, the size of the sample was about one-fourth of what was originally intended. People who respond to surveys are different from people who don't, not only in the obvious way (their attitude toward surveys) but also in more subtle and significant ways. When the response rate is low (as it was in this case, 0.24), a survey is said to suffer from *nonresponse bias*. This is a special type of selection bias where reluctant and nonresponsive people are excluded from the sample.

Dealing with nonresponse bias presents its own set of difficulties. We can't force people to participate in a survey, and paying them is hardly ever a solution since it can introduce other forms of bias. There are ways, however, of minimizing nonresponse bias. For example, the *Literary Digest* survey was conducted by mail. This approach is the most likely to magnify nonresponse bias because people often consider a mailed questionnaire just another form of junk mail. Of course, considering the size of the mailing list, the *Literary Digest* really had no other choice. Here again is an illustration of how a big sample size can be more of a liability than an asset.

Nowadays, almost all legitimate public opinion polls are conducted either by telephone or by personal interviews. Telephone polling is subject to slightly more nonresponse bias than personal interviews, but it is considerably cheaper. Even today, however, a significant segment of the population has no telephone in their homes (in fact, a significant segment of the population has no homes), so that selection bias can still be a problem in telephone surveys.

The most extreme form of nonresponse bias occurs when the sample consists only of those individuals who step forward and actually "volunteer" to be in the sample. A blatant example of this is the 900-number telephone polls, in which an individual not only has to step forward, but he or she actually has to pay to do so. It goes without saying that people who are willing to pay to express their opinions are hardly representative of the general public and that information collected from such polls should be considered suspect at best.

Two morals of the story:

- A badly chosen big sample is much worse than a well-chosen small sample
- Watch out for selection bias and nonresponse bias.