Implementing a Random Sampling Strategy

In this example we are sampling baseball players who have previously been assigned numbers from 1 to 35.

1. Use the table of random numbers in a predetermined order

2. Select pairs of digits, discarding 00 or 36–99, as well as any pair that duplicates a previous pair

3. Choose players until sample is complete.
Players 29, 22, 06, 23, 18, 25, 19, 20, 10, 03, 27, 24, 21, 30, 32, 34

For SRS: Choose first 10 of these.

For stratified sample (4 from P=1 to 18; 1 from C=19 or 20; 3 from I=21 to 28; 2 from O=29 to 35)

P — 06, 18, 10, 03
C — 19
I — 22, 23, 25
O — 29, 30
Sources of bias in surveys

“Bias” refers to a situation where a survey is systematically getting the wrong answer, usually because the survey is not sampling the population correctly.

Numerous sources, such as:

- Undercoverage
- Nonresponse bias
- Response bias
Undercoverage occurs when some part of the population is not covered in the sample.

Nonresponse bias means bias when some of the people don’t respond to the survey.

Response bias isn’t just the opposite of nonresponse bias! It refers to situations where the wording of a question may influence people’s response.

Example —

“Does it seem possible or does it seem impossible to you that the Nazi extermination of the Jews never happened?”
Nonresponse bias

All surveys have some problem with nonresponse. However it’s more problematic if the proportion of nonresponses is large, or if there is some systematic reason to believe that people who would respond one way to a question are more likely to be nonrespondents than people who would respond the other way. Examples:

1. Sometime it’s suggested that supporters on one political party are more likely to refuse to answer the question that supporters of the other

2. Questions about activities of questionable ethics or legality

3. Shere Hite example from text — in a survey of women who had been married at least 5 years, 70% said they had extramarital affairs. But the survey was based on only 4500 respondents out of 100000 people mailed questionnaires.
Other issues in survey sampling: *convenience samples* and *volunteer samples* can be highly problematic.

Example of the latter: Ann Landers survey about whether parents are glad they had children.
Bill Clinton’s Economic Plan

In 1993, shortly after Bill Clinton became President, a California TV station asked viewers to respond to the question “Do you support the President’s economic plan?” Shortly afterwards, a properly conducted opinion survey asked the same question. Here are the two responses:

<table>
<thead>
<tr>
<th></th>
<th>Television Poll</th>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes (support plan)</td>
<td>42%</td>
<td>75%</td>
</tr>
<tr>
<td>No (don’t support plan)</td>
<td>58%</td>
<td>18%</td>
</tr>
<tr>
<td>Don’t know</td>
<td>0%</td>
<td>7%</td>
</tr>
</tbody>
</table>

Moral: the respondents to the TV poll were a self-selected sample (Voluntary Response Bias)
Responses to a Survey About Voting

In 1986, a survey organization asked people whether they had voted in the recent congressional elections. Unknown to the respondents, the survey organization already knew the answers, because they had obtained voting records.

Of the 211 people who had voted, 203 said they voted. But of the 98 who had not voted, 39 claimed that they did.
**Should Congress Repeal the Public Affairs Act?**

A *Washington Post* survey conducted in 1995 asked 1,000 respondents whether the 1975 Public Affairs Act should be repealed. There was no such act! But 43% expressed an opinion about it, with 24% saying it should be repealed and 19% saying it should not.

They then tried a variant: “President Clinton believes the 1975 Public Affairs Act should be repealed. Do you agree?” 36% of those who claimed to be Democrats agreed, but only 16% of those who claimed to be Republicans.

However when the question was changed to: “Republicans in Congress believe the 1975 Public Affairs Act should be repealed. Do you agree?” the percentages were almost exactly reversed.
The Literary Digest poll of 1936

The granddaddy of all biased surveys...

In 1936 a magazine called Literary Digest conducted a poll to ask people who they supported between the two presidential candidates that year, Alf Landon (R) and Franklin Roosevelt (D). They constructed a sampling frame from telephone directories, country club memberships and automobile registrations. They mailed 10 million questionnaires and got 2.3 million responses. They showed Landon would get 57% of the vote. However in the actual election, Landon got only 36% and Roosevelt won in a landslide.
Two problems:

- *Undercoverage* — those who received surveys were not a random sample of the population

- *Nonresponse* — only 23% responded — probably those who were angry with Roosevelt and wanted a change of president were more likely to respond.

Meanwhile George Gallup conducted his own survey of 50,000 people and predicted the correct result.

Moral: Even an enormously large sample can give the wrong answer if the survey is systematically biased.
More recent examples of incorrect opinion polls —

- “Dewey beats Truman” in 1948 (faulty use of quota sampling - p. 188, question 4.85)
- The British general election of 1992
- Al Gore’s faulty concession in 2000
- The New Hampshire Primary of 2008