

## Statistics Is Sexy

CareerCast.com is a job Web site, and its organizers analyzed professions using five criteria: environment, income, employment



prospects, physical demands, and stress. Based on that study,

here are the top ten

jobs: (1) mathematician, (2) actuary, (3) **statistician** (author's emphasis), (4) biologist, (5) software engineer, (6) computer system analyst, (7) historian, (8) sociologist, (9) industrial designer, (10) accountant. Lumberjacks are at the bottom of the list with very low pay, dangerous work, and poor employment prospects.

Reporter Steve Lohr wrote the article "For Today's Graduate, Just One Word: Statistics" in the *New York Times*. In that article he quoted the chief economist at Google as saying that "the sexy job in the next 10 years will be statisticians. And I'm not kidding."

elicit a desired response. Here are the actual rates of "yes" responses for the two different wordings of a question:

97% yes: "Should the President have the line item veto to eliminate waste?"

57% yes: "Should the President have the line item veto, or not?"

**Order of Questions** Sometimes survey questions are unintentionally loaded by such factors as the order of the items being considered. See the following two questions from a poll conducted in Germany, along with the very different response rates:

- "Would you say that traffic contributes more or less to air pollution than industry?" (45% blamed traffic; 27% blamed industry.)
- "Would you say that industry contributes more or less to air pollution than traffic?" (24% blamed traffic; 57% blamed industry.)

**Nonresponse** A *nonresponse* occurs when someone either refuses to respond to a survey question or is unavailable. When people are asked survey questions, some firmly refuse to answer. The refusal rate has been growing in recent years, partly because many persistent telemarketers try to sell goods or services by beginning with a sales pitch that initially sounds like it is part of an opinion poll. (This "selling under the guise" of a poll is now called *sugging*.) In *Lies, Damn Lies, and Statistics*, author Michael Wheeler makes this very important observation:

People who refuse to talk to pollsters are likely to be different from those who do not. Some may be fearful of strangers and others jealous of their privacy, but their refusal to talk demonstrates that their view of the world around them is markedly different from that of those people who will let poll-takers into their homes.

**Missing Data** Results can sometimes be dramatically affected by missing data. Sometimes sample data values are missing because of random factors (such as subjects dropping out of a study for reasons unrelated to the study), but some data are missing because of special factors, such as the tendency of people with low incomes to be less likely to report their incomes. It is well known that the U.S. Census suffers from missing people, and the missing people are often from the homeless or low income groups.

**Precise Numbers** Example 1 in Section 1-1 included a statement that there are 241,472,385 adults in the United States. Because that figure is very precise, many people incorrectly assume that it is also *accurate*. In this case, that number is an estimate, and it would be better to state that the number of adults in the United States is about 240 million.

**Percentages** Some studies cite misleading or unclear percentages. Keep in mind that 100% of some quantity is *all* of it, but if there are references made to percentages that exceed 100%, such references are often not justified. In referring to lost baggage, Continental Airlines ran ads claiming that this was "an area where we've already improved 100% in the last six months." In an editorial criticizing this statistic, the *New York Times* correctly interpreted the 100% improvement to mean that no baggage is now being lost—an accomplishment that was not achieved by Continental Airlines.

The following list identifies some key principles to apply when dealing with percentages. These principles all use the basic notion that % or "percent" really means "divided by 100." The first principle is used often in this book.