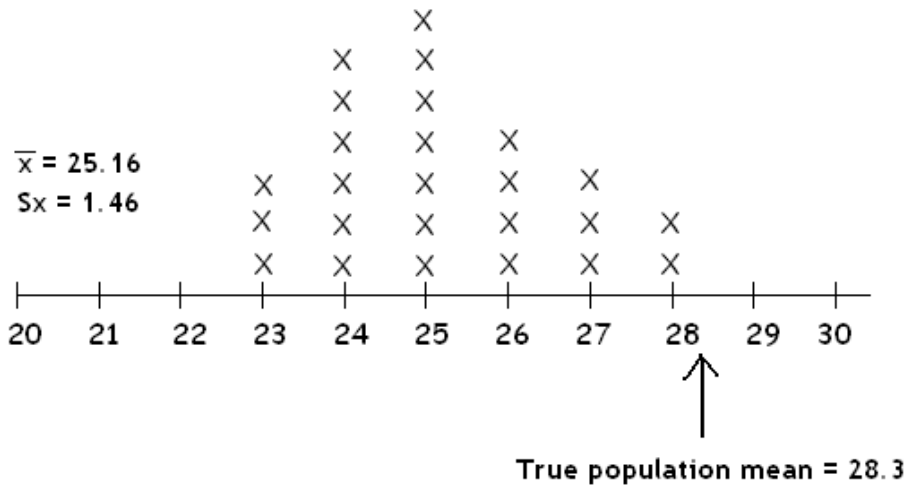


## Statistical Studies: Sources of Variability

### III.C Student Activity Sheet 9: Introduction to Statistical Bias and Variability

Recall Spud Potato Chips one last time. One sample that was “collected” had the results shown below. If the true mean of *all* Spud’s bags is really 28.3 grams, as claimed by the company, this sample appears to have problems. The sample mean is far below the true mean and, in fact, all the bags are below the true mean.



This discrepancy between a population parameter and a sample statistic is known as **statistical bias**, which can result from many different sources. Two broad categories of statistical bias are **biased sampling method** and **biased statistic**.

1. Give an example of a biased sampling method.
2. Sometimes researchers spend vast resources (time, money, effort) to get a great sample and then still end up with a biased statistic. For example, something could be wrong with the data collection method. What could happen after selecting the sample (for example, potato chip bags) and before calculating the statistic that could result in a biased statistic?

## Statistical Studies: Sources of Variability

### III.C Student Activity Sheet 9: Introduction to Statistical Bias and Variability

---

If in fact the method of sampling has introduced bias, one reason for this might be that the sample was actually nonrepresentative of the population (*nonrepresentative sampling*).

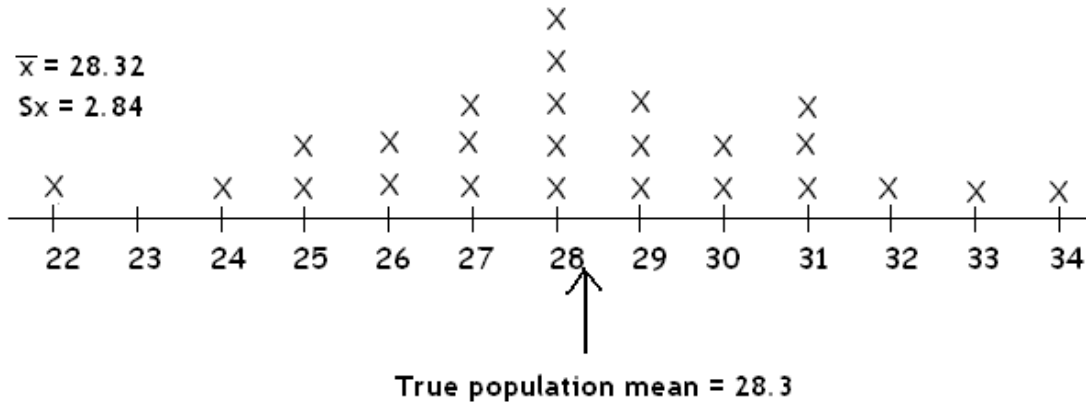
Nonrepresentative sampling is when the sample does not represent the differences in a population. For example, the students in this class are mainly seniors in high school; therefore, they do not represent all students in the school. If only the football team or only the volleyball team were surveyed at most schools, the sample would have all males or all females.

3. What other situations could be responsible for bias? Provide examples where possible.
4. **REFLECTION:** Describe a scenario in which different types of sampling methods lead to different kinds of bias in sampling, including yielding nonrepresentative samples or samples with undercoverage of a population.
5. Biased statistics can result from a variety of problems during the data collection process. Record your thoughts or information you have located through research regarding each of the following sources of statistical bias.
  - a. Response bias
  - b. Nonresponse bias
  - c. Observer effect
  - d. Wording of questions
  - e. Placebo effect

### Statistical Studies: Sources of Variability

#### III.C Student Activity Sheet 9: Introduction to Statistical Bias and Variability

Recall that a second sample of Spud Potato Chips was “collected,” and the following results were obtained:



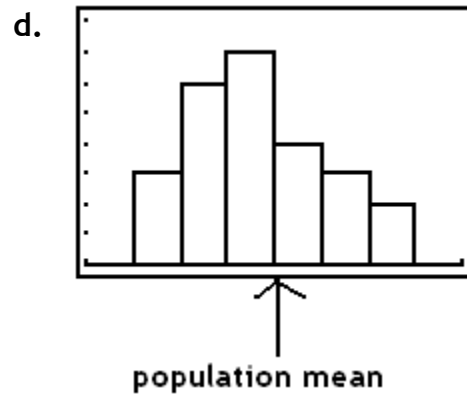
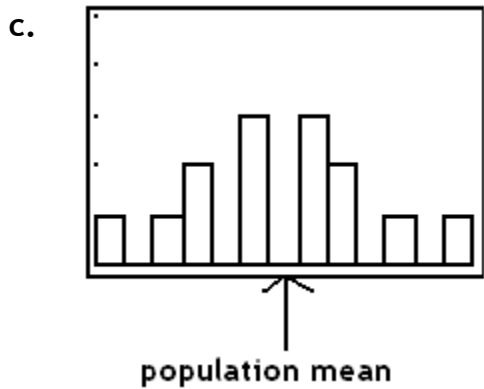
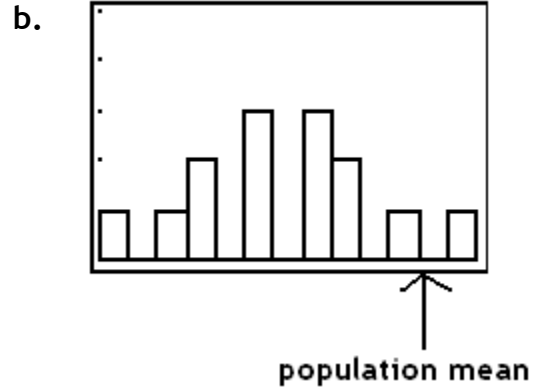
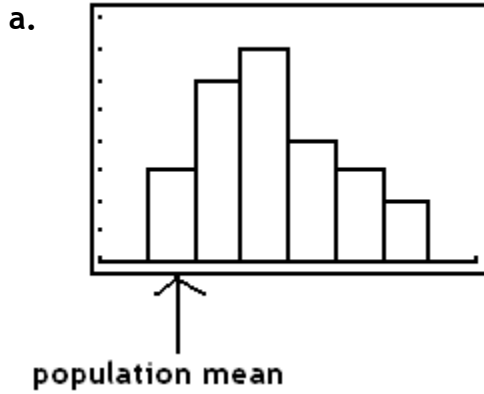
- Comment on this distribution compared to that of the original Spud’s sample.
- REFLECTION:** This sample of chips is an example of *high variability* and no statistical bias. This situation can be a big problem for the manufacturer. People who get bags with only 22 grams of chips probably do not really care what the mean is—they just care that they were shorted! Compare the standard deviations that are given for the original distribution on Page 1 and this distribution.

What are your recommendations for the manufacturer?

### Statistical Studies: Sources of Variability

#### III.C Student Activity Sheet 9: Introduction to Statistical Bias and Variability

8. Identify each of the following as *high* or *low statistical bias* and *high* or *low variability*.



9. **REFLECTION:** In your opinion, which graph in Question 8 represents the worst situation for researchers? Explain your reasoning.

10. **EXTENSION:** Variability can be caused by *natural variability* or *induced variability*. Research each, record your findings, and provide some examples.

- a. Natural variability
- b. Induced variability