1-10. **What did they do? For the following reports about statistical studies, identify the following items (if possible). If you can't tell, then say so—this often happens when we read about a survey.**

a) The population
b) The population parameter of interest
c) The sampling frame
d) The sample
e) The sampling method, including whether or not randomization was employed
f) Any potential sources of bias you can detect and any problems you see in generalizing to the population of interest

1. A business magazine mailed a questionnaire to the human resource directors of all of the Fortune 500 companies, and received responses from 23% of them. Those responding reported that they did not find that such surveys intruded significantly on their workday.

2. A question posted on the Lycos Web site on 18 June 2000 asked visitors to the site to say whether they thought that marijuana should be legally available for medicinal purposes.

3. Consumers Union asked all subscribers whether they had used alternative medical treatments and, if so, whether they had benefited from them. For almost all of the treatments, approximately 20% of those responding reported cures or substantial improvement, in their condition.

4. The Gallup Poll interviewed 1423 randomly selected American citizens September 10-14, 1999 and reported that when "asked which type of content bothers them most on TV, 44% of Americans identify 'violence,' 23% choose 'lewd and profane language/ while 22% say 'sexual situations.'"

5. Researchers waited outside a bar they had randomly selected from a list of such establishments. They stopped every 10th person who came out of the bar and asked whether he or she thought drinking and driving was a serious problem.

6. Hoping to learn what issues may resonate with voters in the coming election; the campaign director for a mayoral candidate selects one block from each of the city's election districts. Staff members go there and interview all the residents they can find.

7. The Environmental Protection Agency took soil samples at 16 locations near a former industrial waste dump and checked each for evidence of toxic chemicals. They found no elevated levels of any harmful substances.

8. State police set up a roadblock to check cars for up-to-date registration, insurance, and safety inspections. They usually find problems with about 10% of the cars they stop.

9. A company packaging snack foods maintains quality control by randomly selecting 10 cases from each day's production and weighing the bags. Then they open one bag from each case and inspect the contents.

10. Dairy inspectors visit farms unannounced and take samples of the milk to test for contamination. If the milk is found to contain dirt, antibiotics, or other foreign matter, the milk will be destroyed and the farm re-inspected until purity is restored.

11. **Parent opinion, part 1.** In a large city school system with 20 elementary schools, the school board is considering the adoption of a new policy that would require elementary students to pass a test in order to be promoted to the next grade. The PTA wants to find out whether parents agree with this plan. Listed below are some of the ideas proposed for gathering data. For each, indicate what kind of sampling strategy is involved and what (if any) biases might result.

   a) Put a big ad in the newspaper asking people to log their opinions on the PTA Web site.
   b) Randomly select one of the elementary schools and contact every parent by phone.
   c) Send a survey home with every student, and ask parents to fill it out and return it the next day.
   d) Randomly select 20 parents from each elementary school. Send them a survey, and follow up with a phone call if they do not return the survey within a week.

12. **Parent opinion, part 2.** Let's revisit the school system described in Exercise 11. Four new sampling strategies have been proposed to help the PTA determine whether parents favor requiring elementary students to pass a test in order to be promoted to the next grade. For each, indicate what kind of sampling strategy is involved and what (if any) biases might result.

   a) Run a poll on the local TV news, asking people to dial one of two phone numbers to indicate whether they favor or oppose the plan.
   b) Hold a PTA meeting at each of the 20 elementary schools and tally the opinions expressed by those who attend the meetings.
   c) Randomly select one class at each elementary school and contact each of those parents.
   d) Go through the district's enrollment records, selecting every 40th parent. PTA volunteers will go to those homes to interview the people chosen.

13. **Wording the survey.** Two members of the PTA committee in Exercises 11 and 12 have proposed different questions to ask in seeking parents' opinions.

   **Question 1:** Should elementary school-age children have to pass high stakes tests in order to remain with their classmates?
   **Question 2:** Should schools and students be held accountable for meeting yearly learning goals by testing students before they advance to the next grade?

   a) Do you think responses to these two questions might differ? How? What kind of bias is this?
b) Propose a question with more neutral wording that might better assess parental opinion.

14. Survey questions. Examine each of the following questions for possible bias. If you think the question is biased, indicate how and propose a better question.
a) Should companies that pollute the environment be compelled to pay the costs of cleanup?
b) Given that 18-year-olds are old enough to vote and to serve in the military, is it fair to set the drinking age at 21?
c) Do you think high school students should be required to wear uniforms?
d) Given humanity's great tradition of exploration, do you favor continued funding for space flights?

d) When the Environmental Protection Agency (EPA) checks a car like mine to predict its fuel economy, what parameter is it trying to estimate?

18. Fuel economy. Occasionally when I fill my car with gas, I figure out how many miles per gallon my car got. I wrote down those results after 6 fill-ups in the past few months. Overall, it appears my car gets 28.8 miles per gallon.
a) What statistic have I calculated?
b) What is the parameter I'm trying to estimate?
c) How might my results be biased?
d) When the Environmental Protection Agency (EPA) checks a car like mine to predict its fuel economy, what parameter is it trying to estimate?

19. Accounting. Between quarterly audits, a company likes to check on its accounting procedures to address any problems before they become serious. The accounting staff processes payments on about 120 orders each day. The next day the supervisor rechecks 10 of the transactions to be sure they were processed properly.
a) Propose a sampling strategy for the supervisor.
b) How would you modify that strategy if the company makes both wholesale and retail sales, requiring different bookkeeping procedures?

20. Happy workers? A manufacturing company employs 14 project managers, 48 foremen, and 377 laborers. In an effort to keep informed about any possible sources of employee discontent, management wants to conduct job satisfaction interviews with a sample of employees every month.
a) Do you see any danger of bias in the company's plan? Explain.
b) Propose a sampling strategy that uses a simple random sample.
c) Why do you think a simple random sample might not provide the representative opinion the company seeks?
d) Propose a better sampling strategy.
e) Listed below are the last names of the project managers. Use random numbers to select two people to be interviewed. Be sure to explain your method carefully.

<table>
<thead>
<tr>
<th>Barrett</th>
<th>Bowman</th>
<th>Chen</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeLara</td>
<td>DeRoos</td>
<td>Grigorov</td>
</tr>
<tr>
<td>Maceli</td>
<td>Mulvaney</td>
<td>Pagliarulo</td>
</tr>
<tr>
<td>Rosica</td>
<td>Smithson</td>
<td>Tadros</td>
</tr>
<tr>
<td>Williams</td>
<td>Yamamoto</td>
<td></td>
</tr>
</tbody>
</table>

21. Quality control. Sammy's Salsa, a small local company, produces 20 cases of salsa a day. Each case contains 12 jars, and is imprinted with a code indicating the date and batch number. To help maintain consistency, at the end of each day Sammy selects three bottles of salsa, weighs the contents, and tastes the product. Help Sammy select the sample jars. Today's cases are coded 07N61 through 07N80.
a) Carefully explain your sampling strategy.
b) Show how to use random numbers to pick the three jars for testing.

22. A fish story. Concerned about reports of discolored scales on fish caught downstream from a newly sited chemical plant, scientists set up a field station in a shoreline public park. For one week they asked fishermen there to bring any fish they caught to the field station for a brief inspection. At the end of the week, the scientists said that 18% of the 234 fish that were submitted for inspection displayed the discoloration. From this information, can the researchers estimate what proportion of fish in the river have discolored scales? Explain.
1. a) Population—Human resources directors of Fortune 500 companies.
b) Parameter—Proportion who don't feel surveys intruded on their work day.
c) Sampling Frame—list of HR directors at Fortune 500 companies.
d) Sample—23% who responded.
e) Method—Questionnaire mailed to all (nonrandom).
f) Bias—Nonresponse. Hard to generalize because who responds is related to the question itself.

b) Parameter—Proportion who feel marijuana should be legalized for medicinal purposes.
c) Sampling Frame—None given; potentially all people with access to Web site.
d) Sample—Those visiting the Web site who responded.
e) Method—Voluntary response (no randomization employed).
f) Bias—Voluntary response sample. Those who visit the Web site and respond may be predisposed to a particular answer.

3. a) Population—All U.S. adults.
b) Parameter—Proportion who have used and benefited from alternative medicine.
c) Sampling Frame—all Consumer's Union subscribers.
d) Sample—Those who responded (random).
e) Method—Questionnaire to all (nonrandom).
f) Bias—Nonresponse. Those who respond may have strong feelings one way or another.

b) Parameter—Type of content that bothers most on TV.
c) Sampling Frame—All U.S. adults.
d) Sample—1423 randomly selected citizens.
e) Method—Not clear.
f) Bias—probably not biased. Conclusions could be generalized.

5. a) Population—Adults.
b) Parameter—Proportion who think drinking and driving is a serious problem.
c) Sampling Frame—Bar patrons.
d) Sample—Every 10th person leaving the bar.
e) Method—Systematic sampling.
f) Bias—Those interviewed had just left a bar. They probably think drinking and driving is less of a problem than do adults in general.

6. a) Population—City voters.
b) Parameter—Not clear; percentages of voters favoring various issues?
c) Sampling Frame—All city residents.
d) Sample—Cluster sample; one block from each district.
e) Method—Convenience sample within each cluster.
f) Bias—Parameter(s) of interest not clear. Sampling within clusters is not random and may bias results.

7. a) Population—Soil around a former waste dump.
b) Parameter—Concentrations of toxic chemicals.
c) Sampling Frame—Accessible soil around the dump.
d) Sample—16 soil samples.
e) Method—Not clear.
f) Bias—Don't know if soil samples were randomly chosen. If not, may be biased toward more or less polluted soil.

8. a) Population—Cars.
b) Parameter—Proportion with up-to-date registration, insurance, and safety inspections.
c) Sampling Frame—All cars on that road.
d) Sample—Those actually stopped by roadblock.
e) Method—Cluster sample of location; census within cluster.
f) Bias—Time of day and location may not be representative of all cars. Other than that, should be able to estimate proportion of all cars.

b) Parameter—Weight of bags, proportion passing inspection.
c) Sampling Frame—All bags produced each day.
d) Sample—Bags in 10 randomly selected cases, 1 bag from each case for inspection.
e) Method—Multistage sampling.
f) Bias—Should be unbiased.

b) Parameter—Proportion passing inspection.
c) Sampling Frame—All dairy farms?
d) Sample—Not clear. Perhaps a random sample of farms, and then milk within each farm?
e) Method—Multistage sampling.
f) Bias—Should be unbiased if farms and milk at each farm are randomly selected.

11. a) Voluntary response. Only those who both see the ad and feel strongly enough will respond.
b) Cluster sampling. One school may not be typical of all.
c) Attempted census. Will have nonresponse bias.
d) Stratified sampling with follow-up. Should be unbiased.

12. a) Voluntary response. Only those who see the show and feel strongly will call.
b) Possibly more representative than part a, but only strongly motivated parents go to PTA meetings.
c) Multistage sampling with cluster sample within each school. Probably a good design if most of the parents in the class respond.
d) Systematic sampling. Probably a reasonable design.

13. a) Answers will definitely differ. Question 1 will probably get many "No" answers, while Question 2 will get many "Yes" answers. This is wording bias.
b) "Do you think standardized tests are appropriate for deciding whether a student should be promoted to the next grade?" (Other answers will vary.)

14. a) Biased toward yes because of "pollute." "Should companies be responsible for any costs of environmental clean-up?"
b) Biased towards no because of "old enough to serve in the military." "Do you think the drinking age should be lowered from 21?"

c) Seems unbiased.

d) Biased toward yes because of "great tradition." Better to ask, "Do you favor continued funding for the space program?"

15. a) Not everyone has an equal chance. People with unlisted numbers, people without phones, and those at work cannot be reached.
   b) Generate random numbers and call at random times.
   c) Under the original plan, those families in which one person stays home are more likely to be included. Under the second plan, many more are included. People without phones are still excluded.

d) It improves the chance of selected households being included.

e) This takes care of phone numbers. Time of day may be an issue. People without phones are still excluded.

16. Cell phones are more likely to be used by higher income individuals. This will cause an undercoverage bias. As cell phones grow in use, this problem will be lessened.

17. a) Answers will vary.
   b) Your own arm length. Parameter is your own arm length; population is all possible measurements of it.
   c) Population is now the arm lengths of your friends. The average estimates the mean of these lengths.
   d) Probably not. Friends are likely to be of the same age, and not very diverse, or representative of the larger population.

18. a) Mean gas mileage for the last six fill-ups.
   b) Mean gas mileage for the vehicle.
   c) Conditions of late may not be typical.
   d) Mean gas mileage for all cars of this make and model.

19. a) Assign numbers 001 to 120 to each order. Use random numbers to select 10 transactions to examine.
   b) Sample proportionately within each type. (Do a stratified random sample.)

20. a) Most likely that all laborers are selected, no managers, and few foremen. Bias may be introduced because the company itself is conducting the survey.
   b) Assign a number from 001 to 439 to each employee. Use a random number table or software to select the sample.
   c) Still heavily favors the laborers.
   d) Stratify by job type (proportionately to the members in each).

21. a) Select three cases at random, then select one jar randomly from each case.
   b) Use random numbers to choose 3 cases from numbers 61 through 80; then use random numbers between 1 and 12 to select the jar from each case.

22. What conclusions they may be able to make will depend on whether fish with discolored scales are equally likely to be caught as those without. It also depends on the level of compliance by fishermen. If fish are not equally likely to be caught, or fishermen are more disposed to bring discolored fish, the results will be biased.