

## 2 Chapter 1

5. That healthier babies are born to mothers who eat lobsters doesn't mean that eating lobster caused the babies to be healthier. Mothers who eat lobster are probably more affluent than the general population and would tend to eat better, be more knowledgeable about proper pre-natal care, have better health care, etc.
7. Motorcyclists that died in crashes in which helmets may have saved their lives could not be present to testify.
9. There are several possible answers. (1) Since tallness is perceived to be a favorable attribute, people tend to overstate their heights; at the very least, people would tend to round to the next highest inch and not to the nearest inch. (2) Many people do not really accurately know their height. (3) Because Americans tend to express height in feet and inches, errors might occur either in converting heights to all inches or in misstatements like 52" for 5'2". (4) Because many cultures express height in centimeters, some people might not know or be able to readily calculate their heights in inches.
11. No. Since the second 5% price cut would be based on a lower price, two consecutive 5% price cuts yield a smaller price reduction than a single 10% price cut. Mathematically, the two consecutive 5% cuts yield a reduction of  $.05x + .05(x-.05x) = .0975x$ , or a 9.75% price cut.
13. Assuming that each of the 20 individual subjects is ultimately counted as a success or not (i.e., that there are no "dropouts" or "partial successes"), the success rates in fraction form must be one of 0/20, 1/20, 2/20, ..., 19/20, 20/20. In percentages, these rates are multiples of 5 (0%, 5%, 10%, ..., 95%, 100%), and values such as 53% and 58% are not mathematical possibilities.
15. a. Since 100% is the totality of whatever is being measured, removing 100% of some quantity means that none of it is left.  
b. Reducing plaque by over 300% would mean removing three times as much plaque as is there, and then removing even more!

### 1-4 Design of Experiments

1. Observational study, since specific characteristics are measured on unmodified subjects.
3. Experiment, since the effect of an applied treatment is measured.
5. Random, since each 212 area code telephone number has an equal chance of being selected. But this is a really a complex situation, as indicated by the following NOTES.  
NOTE 1: This ignores the fact that some residences may have more than one phone number. A residence with two different phone numbers (e.g., one for the parents and one for the teenagers) has twice the chance of being selected as does a residence with a single phone number.  
NOTE 2: The scenario stated the organization sought to poll "residents" with the 212 area code. If the organization polls all residents at each selected number, this is cluster sampling. If the organization polls one resident at each selected number, the sample is not a random sample of "residents" because a resident living alone and having his own phone number has a higher chance of being selected than a resident living with others (e.g., in a

family) and sharing a common phone number.

NOTE 3. The poll will not include residents in the 212 area code who do not have such phone numbers. This is not a problem if the intended population is phone customers (e.g., for a poll of satisfaction with phone service), but it is if the intended population is general residents (e.g., for a poll of satisfaction with garbage service).

7. Convenience, since the sample was simply those who happen to pass by.
9. Stratified, since the population of interest (assumed to be all car owners) was divided into 5 subpopulations from which the actual sampling was done.
11. Cluster, since the population of interest (assumed to be all students at The College of Newport) was divided into classes which were randomly selected in order to interview all the students in each selected class.  
NOTE: Ideally the division into classes should place each student into one and only one class (e.g., if every student must take exactly one PE class each semester, select the PE classes at random. In practice such divisions are often made in ways that place some students in none of the classes (e.g., by selecting from all 2 pm W-W-F classes) or in more than one of the classes (e.g., by selecting from all the classes offered in the college). With careful handling, imperfect divisions do not significantly affect the results.
13. Stratified, since the population of interest (assumed to be all workers) was divided into 3 subpopulations from which the actual sampling was done.
15. Systematic, since every fifth element in the population (assumed to be all drivers passing the checkpoint during its operation) was sampled.
17. There are several possible answers. (1) Write each full-time student's name on a slip of paper, place the slips in a box, mix them thoroughly, and select 200 of them. (2) Assign each full-time a number (e.g., alphabetically), and use a table of random digits (or a calculator or a computer) to generate 200 random numbers with the appropriate numbers of digits.
19. Obtain from each college bookstore a list of the textbooks currently being used, and compile a single master list (i.e., without duplications). Number the textbooks on the master list, and use a table of random digits (or a calculator or a computer) to generate 200 numbers with the appropriate numbers of digits.
21.
  - a. Open questions elicit the respondent's true feelings without putting words or ideas into his mind. In addition, open questions might produce responses the pollster failed to consider. Unfortunately open questions sometimes produce responses that are rambling, unintelligible or not relevant.
  - b. Closed questions help to focus the respondent and prevent misinterpretation of the question. Sometimes, however, closed questions reflect only the wording and opinions of the pollster and do not allow respondents to express legitimate alternatives.
  - c. Closed questions are easier to analyze because the pollster can control the number of possible responses to each question and can word the responses to establish relationships between the responses and with other questions.
23. Confounding occurs when the researcher is not able to determine which factor (often one planned and one unplanned) produced an observed effect. If a restaurant tries adding an evening buffet for one week and it is the same week a nearby theater happens to show a real blockbuster that attracts unusual crowds to the neighborhood, the restaurant can not

## 4 Chapter 1

know whether its increased business is due to the new buffet or the extra traffic created by the theater.

25. a. Possibly; no. Stratified random sampling can employ either the same sample size for each stratum or different sample sizes for the various strata. It results in a random sample only when the sample size for each stratum is proportional to the size of the stratum -- i.e., the same proportion (and not the same number) of each stratum is selected for the sample. If the strata are all the same size, then use the same sample size for each; if one stratum is half the size of the others, then its sample size should be half the other sample sizes. If one stratum is half the size of the others and the same sample size is used for each of the strata, then an element in the smaller stratum has a larger chance of being selected than an element in a larger strata -- and that violates the definition of a random sample that requires that each element has the same chance of being selected. Stratified sampling can never result in a simple random sample. It guarantees that the total sample will always include elements from each of the strata, and that total samples without any elements from one of the strata can not occur -- and that violates the definition of simple random sampling that requires that each total sample has the same chance of being selected.
- b. Possibly; no. When each element in the population is in one and only cluster, cluster sampling always results in a random sample. The chance that any element is selected is the chance that its cluster is selected; since each cluster has the same chance of being selected, each element has the same chance of being selected -- and that satisfies the definition of a random sample. Cluster sampling can never result in a simple random sample. It guarantees that total samples with elements from each of the clusters can not occur -- and that violates the definition of simple random sampling that requires that each total sample has the same chance of being selected.

### 1-5 Introduction to Excel

1. a. A collection of data organized into rows and columns.  
b. Worksheet.  
c. An Excel data file.  
d. 3.
3. a. Items too big for the allotted cells are accepted, but partially hidden behind adjacent occupied cells to the right. Alpha-numeric items are left justified within each cell, and numerical values are right justified within each cell.  
b. Click on the **save** icon (diskette) on the tools bar and follow the  
c. The printed worksheet is reproduced in the same format it appeared on the screen -- i.e. with the same amount of material showing and same within-cell justification, but the cell lines do not appear.  
d. Click on the lower **X** in the upper right corner.  
e. Click on the **open** icon (opening file folder) on the tools bar.  
f. When the **delete** key is pressed the highlighted cells become empty -- i.e., the items that occupied the cells have been deleted. When the **undo** arrow is clicked, the items reappear.