

Statistics: Concepts and Controversies, 10e
Chapter 2 Solutions

For these exercises, remind students that the population is the group about which information is desired. The population is not necessarily the group represented by the sample.

2.1. This is not a simple random sample. To be a simple random sample, every group of eight students has to have the same probability of selection. Not every possible group of eight students can even be selected. For example, students sitting in the same row can never be selected for the same sample.

2.2. Step 1: Label. For the 13 Canadian provinces and territories, we use labels
01, 02, 03,, 11, 12, 13

Specifically, the list of Canadian provinces and territories with labels attached is

- | | |
|------|---------------------------|
| (01) | Alberta |
| (02) | British Columbia |
| (03) | Manitoba |
| (04) | New Brunswick |
| (05) | Newfoundland and Labrador |
| (06) | Northwest Territories |
| (07) | Nova Scotia |
| (08) | Nunavut |
| (09) | Ontario |
| (10) | Prince Edward Island |
| (11) | Quebec |
| (12) | Saskatchewan |
| (13) | Yukon |

Step 2: Software or table. We used the Research Randomizer and requested that it generate one set of numbers with three numbers per set. We specified the number range as 1 to 13. We requested that each number remain unique and that the numbers be sorted least to greatest. We asked to view the outputted numbers with the markers off. After clicking the “Randomize now!” button, we obtained the digits 1, 5, and 12. (Of course, when you use the Research Randomizer, you will very likely get a different set of three numbers.) The sample is the provinces and territories labeled 01, 05, and 12. These are Alberta, Newfoundland and Labrador, and Saskatchewan.

To use the table of random digits, we might enter Table A at line 116 (any line may be used), which is

14459 26056 31424 80371 65103 62253 50490 61181

The two-digit groups in this line are

14 45 92 60 56 31 42 48 03 71 65 10 36 22 53 50 49 06 11 81

We used only labels 01 to 13, so we ignore all other 2-digit groups. The first three labels between 01 and 13 that we encounter in the table choose our sample. Of the labels in line 116, we ignore labels that are too high (over 13). The sample is the provinces and territories labeled 03, 10, and 06. These are Manitoba, Prince Edward Island, and Northwest Territories.

2.3. (b) It systematically favors certain outcomes.

2.4. (c) A convenience sample.

2.5. (c) I write the names of all the students on similar slips of paper, put the slips of paper in a box, mix them well, and draw 10 slips from the box. The 10 names drawn are my sample.

2.6. (a) An SRS because the final sample was obtained by simple random sampling.

2.7. (b) Angela and Lucinda.

2.8. I would not be convinced that a majority of readers oppose the restaurant and bar. Voluntary response samples generally misrepresent the population. People with the strongest opinions (particularly negative) tend to participate in voluntary response samples.

2.9. (a) 466. (b) The online poll is a voluntary response sample, which is often biased.

2.10. (a) 318,600. (b) The online poll is a voluntary response sample, so even though the sample size is very large, the results are likely to misrepresent the population.

2.11. Voluntary response samples are generally biased. In this case, 76% is probably higher than the true proportion of all adults who think Trump should fire Mueller.

2.12. (a) Descriptions may vary. Possible answers are: “Residents of West Lafayette (and Lafayette),” or “motorists.” (b) The proportion favoring one-way streets is almost certainly larger, because only those with strong (usually negative) opinions would take the time to call in.

2.13. (a) Post signs around the college inviting students to phone in their comments on the parking facilities or regulations. (b) Send interviewers out into the parking lots at a particularly busy time of day; instruct them to randomly question students using the parking lot.

2.14. Call-in polls (especially those that cost to call in), and voluntary response polls in general, tend to attract responses from those who have strong opinions on the subject, and therefore they are often not representative of the population as a whole. On the other hand, there is no reason that the results from a scientific polling organization would over-represent any particular group, so the 56.4% in favor of the law is more reliable.

2.15. Call-in polls (especially those that cost to call in), and voluntary response polls in general, tend to attract responses from those who have strong opinions on the subject, and therefore they are often not representative of the population as a whole. A random sample of size 1000 will ideally be representative of the population as a whole. The 300,000 callers might be an impressive number, but the voluntary response of these callers is not trustworthy.

2.16. Results will vary depending on use of software or Table A. Labeling 00–31 alphabetically (down the columns). Beginning on line 132 of Table A, we choose:

25 Rojo
29 Smalling
20 McTominay
16 Lukaku
31 Valencia

2.17. Results will vary depending on use of software or Table A. Labeling 00–26 alphabetically (down the columns). Beginning on line 112 of Table A, we choose:

04 Byron
11 Froese
19 Pacioretty
02 Armia fi
06 Danault
08 Deslauriers

2.18. (a) Use 4-digit labels 0001 to 5341 (or 0000 to 5340).

(b) Answers will vary depending on use of software or Table A. Beginning on line 107 using labels 0001 to 5341 gives

0094, 3831, 4893, 0720, 2417, 4943, 0906, 3600, 4123, 3468, 1683, 4854, 1979, 3844,
3382, 4697, 3936, 4420, 0676, 5130

2.19. (a) For undergraduates, 300/30,000, or 1%, are chosen, and for graduate students 100/10,000, or 1%, are chosen. So, any student, undergraduate or graduate, has a 1-in-100 chance of being selected. **(b)** Each sample contains exactly 300 undergraduates and 100 graduate students. This is not an SRS because not all samples of size 400 are even possible, let alone equally likely. For example, one could not come up with a sample containing 301 undergraduate students and 99 graduate students.

2.20. (a) Label the students from 00001 to 12304 (or 00000 to 12303). **(b)** Answers will vary depending on use of software or Table A. Taking five digits at a time in Table A, beginning at line 125, we choose 12149, 06565, 00795, 08727, and 09517.

2.21. Results will vary depending on use of software or Table A. Labeling 00–35 alphabetically (down the columns). Beginning on line 112 of Table A, we choose:

04 Eolo
11 Jade Mountain
19 Nihi Sumba

2.22. (a) True. All pairs of digits (there are 100 from 00 to 99) are equally likely. (b) False. Such regularity holds only in the long run. If it were true, you could look at the first 39 digits and know whether or not the 40th was a 4. (c) False. Five random digits have a $1/100000$ chance to be 12345, so this sequence will occasionally occur. The sequence 12345 is no more or less random than 99999, or 82507, or any other five-digit sequence.

2.23. Voluntary response samples sponsored by a TV station in San Antonio, such as Twitter polls, will misrepresent the population by having a disproportionate number of San Antonio fans in the sample.

2.24. If those performing surveys always asked their questions of the person who answered the phone, then those who are more likely to answer the phone would be over-represented in the sample. Such people may be more likely to have some particular personality trait, for example; this could introduce bias into the sample.

2.25. (a) The population is adult Hispanic residents of Denver. The sample is made up of one adult from each responding Hispanic household among the 200 selected addresses. (b) The survey will likely underestimate concern about racial profiling because of reluctance to make negative comments about the police to an officer.

Note: *The sample size is not clear in this situation, because some households may not include Hispanics, and some may decline to participate. We can only say that there will be no more than 200 people in the sample.*

2.26. (a) Random selection seems reasonable in this case, although some students might make a case for giving preferential treatment to some group or groups—for example, seniors. (b) Random selection is not the practice used here. Transplant candidates are assessed based on such factors as how long they have been waiting and likelihood of survival with or without the transplant, not to mention tissue typing. More information can be found on the Internet—for example, at the home page of the United Network for Organ Sharing (www.unos.org). (c) Random selection seems to be the best option if voluntary enlistment is inadequate. Conscientious objectors and student deferments might be considered in discussion of this question.