

## Lab 3 – Sections 1.2, 2.1, 2.2, and 2.3 — Due Wed., 9/4

---

The problems on this page are taken from Section 1.2 in the book titled, "Collecting Data: Planning an Observational Study"

For each of the situations described in questions 1 through 5, state whether the sampling procedure is simple random sampling, stratified random sampling, cluster sampling, systematic sampling, or convenience (or voluntary response) sampling.

1. All first-year students at a university are enrolled in one of 30 sections of a seminar course. To select a sample of freshmen at this university, a researcher selects four sections of the seminar course at random from the 30 sections and all students in the four selected sections are included in the sample.

1. \_\_\_\_\_

2. To obtain a sample of students, faculty, and staff at a university, a researcher randomly selects 50 faculty members from a list of faculty, 100 students from a list of students, and 30 staff members from a list of staff.

2. \_\_\_\_\_

3. A student in this class gets the data for his final project by conducting a facebook poll. One of his group partners gets data by asking coworkers for responses. The focus of the group's study was to find out if less than 5% of people ages 18 to 30 have ever tried narcotics.

3. \_\_\_\_\_

4. To obtain a sample of the seniors at a particular high school, a researcher writes the name of each senior on a slip of paper, places the slips in a box and mixes them, and then selects 10 slips. The students whose names are on the selected slips of paper are included in the sample.

4. \_\_\_\_\_

5. To obtain a sample of those attending a basketball game, a researcher selects the 24th person through the door. Then, every 50th person after that is also included in the sample.

5. \_\_\_\_\_

6. What type of bias is present in the sample described in problem 3? Explain the reasoning for your answer.

**The problems on this page are taken from Section 2.1 in the book titled, "Data Classification"**

**For problems 7 through 14, Classify each variable as numerical or categorical.**

- 7. A Gallup poll asked Americans, "Which subject (Math, English, Art,...), if any, has been the most valuable in your life?" 7. \_\_\_\_\_
- 8. Price of a textbook 8. \_\_\_\_\_
- 9. The different flavors of donuts at Dunkin Donuts 9. \_\_\_\_\_
- 10. Heights of NBA athletes 10. \_\_\_\_\_
- 11. The numbers on the back of NBA athletes' uniforms 11. \_\_\_\_\_
- 12. Concentration of a contaminant (micrograms per cubic centimeter) in a water sample 12. \_\_\_\_\_
- 13. Bank account numbers of each person in this class 13. \_\_\_\_\_
- 14. Zip code of each person's residence. 14. \_\_\_\_\_

**For problems 15 through 20, classify each variable as discrete or continuous.**

- 15. The number of people who buy a coffee from Starbucks today. 15. \_\_\_\_\_
- 16. The volume of water (in cubic feet) of each of the Great Lakes. 16. \_\_\_\_\_
- 17. Number of pages in a book 17. \_\_\_\_\_
- 18. The length of a 1-year-old rattlesnake 18. \_\_\_\_\_
- 19. The amount of time it takes to answer this question. 19. \_\_\_\_\_
- 20. The number of nuclear power plants in the world 20. \_\_\_\_\_

**The problems on this page are taken from Section 2.2 in the book titled, "Displaying Categorical Data: Bar Charts and Comparative Bar Charts"**

21. A questionnaire on political affiliation showed information obtained from 50 respondents. Construct a relative frequency distribution for the data (I = independent, D = democrat, R = republican, L = libertarian) given below.

R D I R R I R D R R R D R R D D R D I D R I D R R  
 R L I D I I D L I R R D R D I D R D I D R L D D I

22. The article “Most Smokers Wish They Could Quit” (Gallup Poll Analyses, November 21, 2002) noted that smokers and nonsmokers perceive the risks of smoking differently. The accompanying relative frequency table summarizes responses regarding the perceived harm of smoking for each of three groups: a sample of 241 smokers, a sample of 261 former smokers, and a sample of 502 nonsmokers. Construct a comparative bar chart for these data. Because the three sample sizes are different, don’t forget to use relative frequencies in constructing the bar chart. Comment on how smokers, former smokers, and nonsmokers differ with respect to perceived risk of smoking.

RESPONSE	FREQUENCY		
	Smokers	Former Smokers	Nonsmokers
Very harmful	145	204	432
Somewhat harmful	72	42	50
Not too harmful	17	10	15
Not at all harmful	7	5	5

The problems on this page are taken from Section 2.3 in the book titled, "Displaying Numerical Data: Dotplots, Stem-and-Leaf Plots, and histograms "

23. The number of calories per serving for selected ready-to-eat cereals is listed here. Construct a frequency distribution and histogram using the following class intervals: 80 to < 110, 110 to < 140, etc.

130 190 140 80 100 120 220 220 110 100 210 130 100 90 210 120 200 120 180 120 190 210  
120 200 130 180 260 270 100 160 190 240 80 120 90 190 200 210 190 180 115 210 110 225  
190 130

24. The amount of protein (in grams) for a variety of fast-food sandwiches is reported here. Construct a frequency distribution and histogram using the following class intervals: 10 to < 20, 20 to < 30, etc.

23 30 20 27 44 26 35 20 29 29 25 15 18 27 19 22 12 26 34 15 27 35 26 43 35 14 24 12 23  
31 40 35 38 57 22 42 24 21 27 33