

Chapter 2 Review – Collecting Data

MDM4U

Jensen

Section 2.2 Characteristics of Data

1) Write the word or phrase that is being described

- a) _____: the entire group of individuals that a study aims to gather information about
- b) _____: part of a population that is examined to gather information
- c) _____: an attempt to gather information about every individual member of the population
- d) _____: a study that considers individuals from different groups at the same time
- e) _____: a study that considers individuals over a long period of time.
- f) _____: A variable that takes numerical values for which it makes sense to find an average.
- g) _____: A variable that places an individual into one of several groups or categories
- h) _____: A quantitative variable that can have an infinite number of values in a given interval. Measurable with all real numbers.
- i) _____: A quantitative variable that can take on only a finite number of values within a given range.

2) Identify each of the following variables as qualitative or quantitative. For each quantitative variable, identify whether it is continuous or discrete.

- a) decibels of loudness
- b) eye colour
- c) carpet texture
- d) number of candies
- e) thickness of a book
- f) volume of a drink
- g) length of a pendulum
- h) time of an airplane's descent
- i) measure on the Richter scale

3) For each observational study described, outline what the population of the study was, what the sample was, the main variables involved (including what type of variables they are), and whether a cross-sectional or longitudinal study was done.

a) A science teacher was interested to find the relationship between homework completion and test performance for grade 10 science students. The teacher chose two grade 10 science students from his class to be part of the study. The teacher checked their homework everyday of the semester and tracked how changes in amount of homework completed affected test performance for each unit.

b) A market researcher wanted to know if teenage males or females were more likely to smoke. She decided to go to a local high school and ask all of the students if they smoked or not.

Section 2.3 Sampling Principles

4) Which type of sampling method is being described?

a) _____: divide the population into groups; randomly select a few of those groups and then sample all members from the selected groups.

b) _____: randomly choose some starting point; then select every n^{th} element in the population, where n is the sampling interval.

c) _____: each member of the population is equally likely to be chosen and the members of the sample are chosen independently of one other

d) _____: the population is organized into groups, a simple random sample of groups is chosen, and then a simple random sample of people within the chosen groups is taken.

e) _____: choosing individuals from the population who are easy to reach

f) _____: the population is divided into groups. A simple random sample of the members of each group is then taken. The size of the sample for each group is proportionate to the group's size

g) _____: a sample that consists of people who choose themselves by responding to a general invitation.

5) Identify the type of random sampling in each of the following scenarios.

a) 80 students are taking an AP stats course and the teacher wants to randomly pick out a sample of 10 students to try out a practice exam. She numbers the students 1 through 80 and uses the random number generator function on her calculator to determine which 10 students will write the practice exam.

b) The Ontario government randomly selects five high schools in Ontario and surveys each teacher in those five schools.

c) Every fiftieth family in the Unionville telephone book is surveyed by phone.

d) To analyze the quality of Canada's health care system, 10% of the hospitals from each province are randomly selected to participate in a study.

e) In order to determine how much sleep typical high school students get, a student surveyed the first 100 students to arrive at school on a particular morning.

f) A magazine posed the question "should drivers be banned using all cell phones?" Readers were encouraged to vote online at the magazines website.

g) There are 30 NBA basketball teams. Each NBA basketball team has 12 players on its roster. To find out if the players think the season is too long, the commissioner decides to randomly select 5 teams, and then randomly select two players from each of those teams to interview.

6) A standard deck of 52 cards has four suits(hearts, diamonds, clubs, and spades), each with 13 cards. If you wanted a random sample of 25% of these cards, describe in detail how you could do this using each of the following methods:

a) simple random sampling

b) systematic random sampling

c) cluster random sampling

7) From a list of 100 grade 12 students numbered 00 to 99, a sample of ten is taken. For each example below, identify what sampling method was used. Justify your choices.

a) 7, 17, 27, 37, 47, 57, 67, 77, 87, 97

b) 30, 31, 32, 33, 34, 35, 36, 37, 38, 39

c) 25, 21, 29, 28, 20, 52, 54, 50, 57, 51

d) 5, 14, 28, 38, 41, 55, 68, 70, 84, 92

8) In your school you have the following demographics:

a) Explain how you could obtain a simple random sample of 20% of the students

Grade	Number of Students
9	120
10	200
11	130
12	190
Total	640

b) Explain how you could obtain a stratified random sample by grade of 20% of the students

c) Explain how you could obtain a systematic random sample of 20% of the students.

Section 2.4 – Survey Design and Bias

9) On the topic of school, create one example for each of the following types of questions:

a. Open question	b. Rating question
c. Ranking question	d. Checklist question

10) Determine whether each example is a(n): information question, checklist question, ranking question, or rating question.

a) Please provide the following information: Name: _____ Age: _____

b) Rank each of the following in the order of importance, where 1 is most important and 6 is least important. It is important to me that my friends are:

_____ honest, _____ trustworthy, _____ friendly, _____ sharing, _____ funny, _____ kind.

c) Which of the following types of music do you enjoy? (Circle as many as apply.)

Classic Rock Jazz Hip-hop Reggae Heavy metal Alternative Country

d) Rate your preference for the following foods (1 means “not at all”, 10 means “very much”)

_____ Ice cream _____ pie _____ lasagna _____ pizza

_____ hot dogs _____ french fries _____ Greek salad _____ soup

11) Identify what type of bias is being indicated.

a) _____: Refers to anything in the survey design that influences the responses. This includes wording of questions, and unwillingness of respondent to reveal personal facts.

b) _____: When the chosen sample does not accurately represent the population

c) _____: When one type of respondent is overrepresented because groupings of different sizes are polled equally instead of proportionately

d) _____: Occurs when an individual chosen for the sample can't be contacted or refuses to participate

12) Identify the type(s) of bias that may result from each of the following data collection methods.

a) You wish to find out how many hours teenagers spend playing video games on an average school night, so you randomly survey five students from each block A class.

b) You wish to determine how many students will come to an upcoming dance and so you send a survey to all Grade 9 classes.

c) Teachers of a school want to gather data on how many students cheat during tests so they conduct one on one interviews with each of their students.

d) The cafeteria hand out surveys during lunch to students to gather information about what types of meals students would like to have. They ask that students complete the surveys during lunch and drop them before class.

e) You are interested in determining how many hours of television teenagers in your school watch per week, so you post a poll on EDSBY and ask for volunteers to fill it out.

13) Consider each of the following questions. Use the criteria for what to avoid in good questions. Identify any criteria (there may be more than one) which are not met and rewrite the question to improve it.

a) On a scale of 1 to 5 (5 being the highest), evaluate the talent level of Sidney Crosby, who was voted by the NHLPA to be the most valuable player in the NHL multiple times.

Criteria not met:

Rewrite:

b) From the list below, indicate which goaltenders you wouldn't choose for your fantasy hockey team.

Roberto Luongo Marc-Andre Fleury James Reimer Martin Brodeur Steve Mason

Jimmy Howard Jonathan Bernier Carey Price Devan Dubnyk Craig Anderson

Criteria not met:

Rewrite:

Section 2.5 Experiment Design

14) Does reducing screen brightness increase battery life in laptop computers? To find out, researchers obtained 30 new laptops of the same brand. They chose 15 of the computers at random and adjusted their screens to the brightest setting. The other 15 laptop screens were left at the default setting – moderate brightness. Researchers then measured how long each machine’s battery lasted. Was this an observational study or an experiment?

15) Does eating dinner with their families improve students’ academic performance? According to an ABC News article, “Teenagers who eat with their families at least five times a week are more likely to get better grades in school. This finding was based on a sample survey conducted by researchers at Columbia University. Was this an observational study or an experiment?

16) A study published in the New England Journal of Medicine (March 11, 2010) compared two medicines to treat head lice: an oral medication called ivermectin and a topical lotion containing malathion. Researchers studied 376 households in areas around the world. Of the 185 households randomly assigned to ivermectin, 171 were free from head lice after 2 weeks compared with only 151 of the 191 households randomly assigned to malathion.

a) Identify the experimental units

b) What are the explanatory and response variables

c) What are the treatments used

17) What are the four principles of experimental design?

a) _____: Use a design that compares two or more treatments

b) _____: Use chance to assign experimental units to different treatments.

c) _____: Keep other variables (besides the ones you are testing) that might affect the response of the subject the same for all groups.

d) _____: Use enough experimental units in each group so that any differences in the effects of the treatments can be distinguished from chance differences between groups

18) Researchers in Canada performed an experiment with university students to examine the effects of in-class laptop use on student learning. All participants in the study were first year undergraduate students enrolled in the business administration program. They were asked to attend a university-style lecture and take notes with their laptops. Half of the participants were assigned to complete other non-lecture related online tasks during the lecture. These tasks were meant to imitate typical student Web browsing during classes. The remaining students simply took notes with their laptops. To assign the treatments, the researchers printed 40 papers with instructions (20 with multitasking and 20 without), shuffled them, and handed them out at random to students in the classroom. At the end of the lecture, all participants took a comprehension test to measure how much they learned from it. The results: students who were assigned to multitask did significantly worse (11%) than students who were not assigned to multitask.

Explain how each of the four principles of experimental design was used in the multitasking study.

19) Nurse practitioners are nurses with advanced qualifications who often act much like primary care physicians. Are they as effective as doctors at treating patients with chronic conditions? An experiment was conducted with 1316 patients who had been diagnosed with asthma, diabetes, or high blood pressure. Within each condition, patients were randomly assigned to either a doctor or a nurse practitioner. The response variables included measures of the patients' health and of their satisfaction with their medical care after 6 months.

a) Which are the blocks in this experiment: the different diagnoses (asthma, and so on) or the type of care (nurse or doctor)?

b) Explain why a randomized block design is preferable to a completely randomized design here.

20) Twenty overweight females have agreed to participate in a study of the effectiveness of four weight-loss treatments: A, B, C, and D. The researcher first calculates how overweight each subject is by comparing the subject's actual weight with her "ideal" weight. The subjects and their excess weights in pounds are as follows:

Birnbaum	35	Hernandez	25	Moses	25	Smith	29
Brown	34	Jackson	33	Nevesky	39	Stall	33
Bunk	30	Kendall	28	Obrach	30	Tran	35
Cruz	34	Loren	32	Rodriguez	30	Wilansky	42
Deng	24	Mann	28	Santiago	27	Williams	22

The response variable is the weight lost after 8 weeks of treatment. Previous studies have shown that the effects of a diet may vary based on a subject's initial weight.

a) Explain why a randomized block design would be better than a completely randomized design in this setting.

b) Should researchers form blocks of size 4 based on subjects' last names in alphabetical order or by how overweight the subjects are? Explain.

c) Explain how you could carry out the random assignment required by your design. Explain your method clearly.

21) In an interesting experiment, researchers examined the effect of ultrasound on birth weight. Pregnant women participating in the study were randomly assigned to one of two groups. The first group of women received an ultrasound; the second group did not. When the subjects' babies were born, their birth weights were recorded. The women who received the ultrasounds had heavier babies.

a) Did the experimental design take the placebo effect into account? Why is this important?

b) Was the experiment double-blind? Why is this important?

c) Based on your answers to a) and b), describe an improved design for this experiment.