

Chapter 1 Textbook exercises

Solutions to even-numbered questions
Statistics and statistical programming
Northwestern University
MTS 525

Aaron Shaw

September 21, 2020

All questions taken from the *OpenIntro Statistics* textbook, 4th edition, Chapter 1.

1.6

- (a) Many possible answers here. A basic one rearranges the first sentence of the question in the textbook: "Is there a difference in unethical behaviors by people from different (perceived) social classes?"
- (b) 129 UC Berkeley undergraduate students.
- (c) From the description in the question text it seems like there are two primary measures:
 - Unethical behavior (candies taken): a discrete numerical measure.
 - Perceived social class: a categorical measure.

1.10

- (a) Each row represents the data collected about a single participant in the survey.
- (b) There were 1,691 participants.
- (c) See the table below:

variable	type	sub-type (if applicable)
sex	categorical	
age	numerical	discrete (rounded to year)
maritalStatus	categorical	
grossIncome	categorical	ordinal
smoke	categorical	
amtWeekends	numerical	discrete
amtWeekdays	numerical	discrete

1.16

- (a) The population of interest is all people. The sample is the 129 UC Berkeley undergraduates who participated in the study.
- (b) Given that this is an observational study conducted on a convenient sample of UC Berkeley undergraduate students any claims to either causal identification or generalizability seem... implausible.

1.40

- (a) The explanatory (independent, predictor, x) variable is percent of a county's population with a bachelor's degree and the response (dependent, outcome, y) variable is each county's per capita income (measured in thousands of US\$).
- (b) There is a positive, linear relationship between the two variables. There are few counties where more than 50% of residents hold bachelor's degrees and few counties with a per capita income greater than \$40k.
- (c) No. Based on the description a causal interpretation is not justified. The data suggest a *positive association* between education and income.

1.42

- (a) This is an observational study.
- (b) The explanatory (independent, predictor, x) variables are child's screen time, sex, and age and mother's education, ethnicity, psychological distress, and employment. If you just said child's screen time that would probably be okay since that seems to be the key explanatory variable that is the focus of the study and the other variables are included to support a more accurate estimate of the relationship between screen time and psychological well-being.
- (c) The response (dependent, outcome, y) variable is child's psychological well-being.
- (d) The best answer to this depends on the target population of the study. The surveys come from three nationally representative samples from the UK, Ireland, and the United States. If the target population of the study is the populations of those three countries, then sure, the study results should generalize. If the target population is "all children on earth" or something like that, well then there's no reason to believe it generalizes since these three countries are in no way representative of the world.
- (e) The study is observational and lacks any clear strategy for identifying a causal relationship between screen time and psychological well-being. As a result, it does not support any claims to have identified causal effects.