Ch.1 Problem 1.
1 Suppose that you are asked to conduct a study to determine whether smaller class sizes lead to improved student performance of fourth graders.
   (i) If you could conduct any experiment you want, what would you do? Be specific.
   (ii) More realistically, suppose you can collect observational data on several thousand fourth graders in a given state. You can obtain the size of their fourth-grade class and a standardized test score taken at the end of fourth grade. Why might you expect a negative correlation between class size and test score?
   (iii) Would a negative correlation necessarily show that smaller class sizes cause better performance? Explain.

Ch.1 Problem 3.
3 Suppose at your university you are asked to find the relationship between weekly hours spent studying \((\text{study})\) and weekly hours spent working \((\text{work})\). Does it make sense to characterize the problem as inferring whether \(\text{study}\) “causes” \(\text{work}\) or \(\text{work}\) “causes” \(\text{study}\)? Explain.

Ch. 1 Problem C3.
C3 The data in MEAP01.RAW are for the state of Michigan in the year 2001. Use these data to answer the following questions.
   (i) Find the largest and smallest values of \(\text{math}4\). Does the range make sense? Explain.
   (ii) How many schools have a perfect pass rate on the math test? What percentage is this of the total sample?
   (iii) How many schools have math pass rates of exactly 50%?
   (iv) Compare the average pass rates for the math and reading scores. Which test is harder to pass?
   (v) Find the correlation between \(\text{math}4\) and \(\text{read}4\). What do you conclude?
   (vi) The variable \(\text{exp}pp\) is expenditure per pupil. Find the average of \(\text{exp}pp\) along with its standard deviation. Would you say there is wide variation in per pupil spending?
   (vii) Suppose School A spends $6,000 per student and School B spends $5,500 per student. By what percentage does School A’s spending exceed School B’s? Compare this to \(100 \cdot \frac{\log(6,000) - \log(5,500)}{\log(6,000)}\), which is the approximation percentage difference based on the difference in the natural logs. (See Section A.4 in Appendix A.)