Show **all** of your work on this assignment and answer each question fully in the given context. *Please* staple your assignment!

1. **Chapter 1, Exercise 1 (page 23)**
2. **Chapter 1, Exercise 9 (page 24)**
3. **Chapter 2, Section 2, Exercise 1 (page 37)**
4. **Lurking variables.**

   There is an common saying in the sciences about establishing cause and effect: "correlation does not imply causation". It summarizes the idea that two events can have a relationship without one causing the other. Examples are easy to think of - suppose I did an observational study in which I looked at the percent of people using an umbrella and the percent of people using windshield wipers, recording the two percentages every hour of the day for two weeks. After collecting my data, I will discover that every time the percent of people using umbrellas rose the percent of people using windshield wipers also rose. The reason for this is not that people using umbrellas cause people to use windshield wipers. Instead both umbrella use and windshield wiper use are responses to a third variable, specifically whether or not it is raining. If I had made note of that during the process, I might have noticed that. In statistics we call variables that are unobserved but potentially the true causative factor **lurking variables** - they variables are not directly seen in the data set, but their influence is.

   Consider the following cases where two variables are related and suggest a possible lurking variable in each case that might explain why the relationship is not a cause and effect. (*note: this question is fairly open ended - as long as you attempt to explain how the lurking variable you suggest could be driving both of the other variables you will get full credit. For example, in the umbrella/windshield wiper example, an satisfactory response would be "The presence of rain may be a lurking variable. When it rains, people use umbrellas to stay dry and windshield wipers in order to drive safely. When it is not raining, neither umbrellas or windshield wipers are needed.*).

   (a) People are more likely to get sick during the winter. They are less likely to get sick during the summer.

   (b) People are more likely to pass away after retiring from work than they are if they are not planning to retire any time soon.

   (c) When lots of swimsuits are being sold, the number of shark attacks is higher. When few people are buying swimsuits, the number of shark attacks is lower.

5. **Hockey game attendance.**

   Caroline performs the following study to see if outside temperature has an effect on attendance at her college's hockey games. For each hockey game at her college, Caroline records the outside temperature and the attendance. Here are her results:

Date		*Temperature, deg. F*	*Attendance*
Friday	12/14	35	840
Wednesday	12/19	20	560
Tuesday	1/ 8	−5	340
Friday	1/11	23	775
Wednesday	1/23	14	680
Saturday	2/ 2	30	950
Friday	2/ 8	28	950

(a) Is this an experiment or observational study?

(b) What type of variable is attendance?

Caroline analyzes her results and finds that outside temperature and attendance have a strong positive correlation (i.e., as one increases, the other also increases). She concludes that higher game day temperatures causes higher attendance at their college's hockey games.

(c) Did she come to a proper conclusion for this study? Why or why not?

(d) Look at the day of the week of the hockey games. What type of variable is this?

(e) How does game attendance relate to the day of the week (regarding weekday games vs. weekend games)?

(f) For what type of studies do you have to worry about possible lurking variables affecting the results?

6. **Washer stretching.**

George works for a company that manufactures rubber washers. He randomly selects 1000 washers off the assembly line throughout two weeks for a study on the durability of these washers under stretching. To make sure that the washers are fit to be used in the real world, George must test the washers. Holding heat constant, George subjects each washer to one of various degrees of stretching. The washers are assigned to the different stretching groups randomly. After each test, George classifies a washer as either defective or non-defective.

(a) Is this an experiment or observational study?

(b) What type of variable is heat?

(c) What type of variable is the amount of stretching?

(d) What type of data is recorded, quantitative or qualitative?

(e) The 100 selected washers constitutes the sample. What is the population?

(f) George analyzes the results and finds that the defect rate increases with the amount of stretching. Can George conclude that the amount of stretching causes a change in the defect rate of the washers? Why or why not?

7. **Microwave popcorn.**

A study is to be carried out to determine the optimal combination of microwave oven settings for microwave popcorn. Cooking time has three possible settings (3, 4, and 5 minutes) and cooking power has two settings (low power, high power). The response (to be minimized) is the number of burned plus the number of unpopped kernels. There are two microwaves available (labeled Microwave A and Microwave B), and due to time constraints both must be used. The popcorn for the experiment are 24 bags of Pop Secret popcorn.

(a) Is this an experiment or an observational study? Explain briefly.

(b) Identify the experimental units.

(c) Identify the three experimental variables and levels of these experimental variables.

(d) Identify a potential blocking factor.

(e) Describe how randomization would be performed in this study if 4 bags of popcorn are assigned to each treatment group.

8. **Measurement characteristics.**

Students Jack and Jill each use the same ruler to measure the thicknesses of various books book to the nearest millimeter. All books have the same thickness of 3.6 cm, a fact that is unknown to them. After doing the measurement five times, the following data are recorded:

Jack: 3.5 cm, 3.7 cm, 4.0 cm, 3.3 cm, 3.5 cm
Jill: 3.9 cm, 3.8 cm, 4.0 cm, 3.9 cm, 3.8 cm

(a) Who has more accurate data?

(b) Who has more precise data?