1. For each of the following situations, write a few sentences describing the population of interest, the inferential objective, and how you might go about collecting a sample. *(from 1.1)*

   (a) A university researcher wants to estimate the proportion of U.S. citizens between the ages of 22 and 35 who are interested in starting their own business.

   (b) A city engineer wants to estimate the average weekly water consumption for single-family dwelling units in the city.

   (c) An electrical engineer wants to determine whether the average length of life of a certain type of transistors is greater than 500 hours.

2. Write a few sentences explaining the difference between a population and a sample, and the logic used to make an inference about a population from a sample.

3. According to the text, the objective of statistics is to make an inference about a population based on a sample, with an associated measure of goodness for this inference. How to quantify this "measure of goodness' is a major theme of 4102. For now, write a few sentences about why a measure of goodness is necessary. You may want to consider what an inference with no measure of goodness might be worth and why.

4. AIDS has become one of the most devastating diseases in modern society (and this was written back in 1992!). The numbers of cases of AIDS (in thousands) reported in 25 major cities in the United States during 1992 are as follows. *(from 1.4)*

38.3	6.2	3.7	2.6	2.1
14.6	5.6	3.7	2.3	2.0
11.9	5.5	3.4	2.2	2.0
6.6	4.6	3.1	2.2	1.9
6.3	4.5	2.7	2.1	1.8

   (a) Construct a relative frequency histogram to describe these data. Don't worry about being too precise, but be neat enough it to be clear to the TA what you did.

   (b) What proportion of these cities reported more than 10,000 cases of AIDS in 1992?

   (c) If one of the cities is selected at random from the 25 for which the preceding data were taken, what is the probability that it will have reported fewer than 3000 cases of AIDS in 1992?

5. For the AIDS data above, calculate the sample mean $\bar{y}$ and the standard deviation s. You may find this easiest by using a spreadsheet program or even specialized statistical

software, if you already know it. Describe the method you used. Now calculate the interval $\bar{y} \pm 2s$. Count the number of measurements that fall within each interval and compare this result with the number that you would expect using the empirical rule.

You probably noticed that there is one very large value (38.3). Eliminate this value and repeat the above calculations. What differences do you see? *(from 1.13 & 1.14)*

6. It has been projected that the average and standard deviation of the amount spent online using the Internet (remember, our text is from 2002) are, respectively, 14 and 17 hours per person, per year. *(from 1.8)*

   (a) What value is exactly one standard deviation below the mean?

   (b) If the amount of time spent online is approximately normally distributed, what proportion of the users spent an amount of time online that is less than the value from part a?

   (c) Is the amount of time spent online approximately normally distributed? Why or why not?

7. Weekly maintenance costs for a factory tend to have an approximately normal distribution with an average of \$420 and a standard deviation of \$30. If \$450 is budgeted for next week, what is an approximate probability that this budgeted figure will be exceeded?

8. Suppose two dice are tossed and the numbers on the upper faces are observed. Let S denote the set of all possible pairs that can be observed. Define the following subsets of S:

   A: The number on the second die is even.

   B: The sum of the two numbers is even.

   C: At least one number in the pair is odd.

   List the points in A, $\bar{C}$, $A \cap B$, $A \cap \bar{B}$, $\bar{A} \cup B$, and $\bar{A} \cap C$. Use notation such as $(2, 3)$ to indicate a two on the first die and a three on the second. *(from 2.4)*

9. From a survey of 60 students attending a university, it was found that 9 were living off campus, 36 were undergraduates, and 3 were undergraduates living off campus. *(from 2.6)*

   (a) Find the number of these students who were undergraduates, were living off campus, or both.

(b) Find the number of these students who were undergraduates living on campus.

(c) Find the number of these students who were graduate students living on campus.

Some of these problems are taken directly from our text.