3.14  
a) The data is experimental data because we are imposing a “treatments” on these subjects (measured by “milk level”) and observing the response to the treatment (TBBMC levels).

b) The control group data is part of the data classified in part a) of this problem: it used to “evaluate the effect of the additional milk in the diet.” Accordingly, the control group data is “experimental.” We have defined experimental data as something that we get after doing something to the experimental units. In this case, you need to exercise caution, as you might be tempted to argue that since nothing was “done” to the control group, their data was not experimental. But it was gathered in a particular context, namely to understand how milk levels and age influence TBBMC. Accordingly, it makes most sense to think of the control group as having received a “zero-level” treatment.

c) The growth rate of TBBMC is a response variable. Age is an explanatory variable. (The study seems to be looking at the effects both of additional milk, and age.)

3.15 Discussed in class

3.16 Discussed in class

3.24 Discussed in class

3.26  
a) A control group helps eliminate the placebo effect.

b) People who take part in the study might be inclined to believe that they are in the hands of experiments, and therefore that whatever they do has beneficial effect, even if in fact it doesn’t.

c) A placebo treatment might be something simple but clearly trivial, like the advice “think of the color blue as you eat your breakfast.” It would be hard to argue that this had any real bearing on health, but to those susceptible to the placebo effect, it might be enough to trigger the effect.

3.27 Discussed in class

3.52  
a) One suspects that the intended “population” is the set of 76 4th year seniors. The population could be the set of all seniors over all time, but since the committee probably collects opinions, and opinions will change as administrative policies change, the 20 seniors chosen this year would not be particularly representative of this broader population. Similarly, one could argue that the population is the set of all undergraduate English majors, but once again, this sample (consisting exclusively of 4th years) would not be representative.

b) The sample is clearly the 20 4th year English majors that were selected.

c) One assumes 4th year students were chosen because they “know the ropes” and are in a good position to provide informed commentary about the undergraduate English major experience. Including 1st year students in the sample would expand the points of view, but the question then becomes: would opinions based on less experience add to the knowledge base? The true answer really depends on the purpose of the committee, and the reason undergraduates were elected to serve on it at all.

3.56  
a) Multi-stage sample. Note that this is slightly different from a stratified random sample, in the sense that first choosing “3 sections” is not dividing the population into “similarity” based strata. (See part d. for a comparison.) Instead, the sampling simply proceeds in stages: first pick sections, then pick an SRS from each section.

b) An SRS of size 5—each group of 5 has the same chance of being chosen.

c) Voluntary response sample—no one is selecting the participants, they select themselves.

d) Stratified random sample—the population is first divided into “gender” strata, and then an SRS is taken in each stratum.
3.58  a) This sample does not necessarily represent the full text: perhaps the 3rd chapter is easy, or hard. A better approach would be to look at a small amount of material from each chapter, or from randomly selected chapters.

b) Only the foolhardy or the manically devoted take a class at 7:30 am. Their perspective is valuable, but maybe not representative of the college at large. A better approach is handing out surveys in all classes, or at least a representative subset.

c) Who knows what subtle bias is inserted into the study this way? Suppose, for example, Armenians tend to have names that begin disproportionately with the letter A; then Armenians will be overrepresented in this study. Much better would be to choose the SRS by assigning everyone a number and using a random number table.

3.62  a) The population seems to be the set of all students at the college.

b) The population seems to be the set of all people living in America, since the study is about all Americans (that is the point of the US census.) The sample is the 250000 people that were contacted by mail, phone, or in person.

c) The population is probably intended to be the set of all people living in the country to which the calls were made, assuming that was a single country. If the calls spanned several countries, presumably the intended population was all people living in these countries.

3.67  Discussed in class

3.69  Discussed in class