Answers to Review Questions and Problems

[PE: This title is from the previous edition. But section title in the chapter text is "Review Problems" not "Review Questions and Problems"]

Chapter 1

1. Science is a systematic and controlled way of gathering information about the world. Methods are integral to science because scientific results are only trustworthy when the procedures used to reach them are considered correct by others in the scientific community.
2. A population is the universe of people, places, or objects that a researcher wishes to study. Populations are usually too large to be studied directly.
3. Samples are subsets of populations. Researchers draw samples because populations are too large to be studied directly. Samples are smaller and therefore more feasible to work with.
4. Replication is important in science because the emphasis on methods means that one person's findings should be able to be replicated by another person using similar methods. It is also important to examine findings using different methods to see if the results remain stable.
5. Hypothesis testing is used to test individual components of theories as a means of determining the validity of those theories. Samples are used to make inferences about populations. Evaluation research assesses the effectiveness of a program or intervention. People, places, or objects are measured before and after an intervention, or a control group is used for comparison to the treatment group. Exploratory research delves into new areas of study about which little is known. Hypotheses are generally not possible, as researchers usually do not have theory or prior evidence to guide them. Descriptive research analyzes a sample and provides basic information about those people, places, events, or objects. No inferences are made to the population.
6. Any three theories and hypotheses derived therefrom; the correctness of students' responses is up to the instructor's judgment.
7. Any three programs or policies; the correctness of students' responses is up to the instructor's judgment.
8. Exploratory research would be used, because the researcher is studying a topic or phenomenon about which very little is known.
9. This would be descriptive research, because that type of research focuses solely upon a sample rather than using a sample to draw conclusions about a population.
10. GIGO stands for *garbage in, garbage out*. In the context of statistics, GIGO means that the results of a statistical analysis are only as accurate and trustworthy as the data and the methods used to collect the data. Statistics cannot fix data gathered using incorrect methods, and incorrect statistics can ruin high-quality data. If the data or the statistics are bad, the results will be wrong.

Chapter 2

1.
a. education

respondents have less education. (This seems counterintuitive, but it is probably a generation effect. Older adults and seniors in the GSS sample might have come from generations where college education was not as prevalent as it is among the younger generations.) The magnitude is very weak, and the coefficient of determination is $(-.078)^2 = .006$, meaning that only .6% of the variance in education is attributable to age. Despite statistical significance, then, these two variables are not related in a substantively meaningful sense.

Chapter 14

Note: Rounding, where applicable, is to two decimal places in each step of calculations and in the final answer. For numbers close to zero, decimals are extended to the first non-zero number. Calculation steps are identical to those in the text; using alternative sequences of steps might result in answers different from those presented here. These differences might or might not alter the final decision regarding the null. Numbers gleaned from SPSS output are presented using three decimal places.

1. Regression's advantage over correlation is its ability to predict values of the DV (y) using specified values of the IV (x). Because regression fits a line to the data, any value of y can be predicted using any value of x. Values of x can be plugged into the formula for the line of best fit ($\hat{y} = a + bx$, at the bivariate level) and the equation can be solved to produce the predicted value of y at the given value of x.

2. The dependent variable must be continuous and normally distributed.

3. No. The OLS model can accommodate IVs of any level of measurement.

4. Multiple regression's advantage over bivariate regression is its ability to account for more than one IV at a time. Throughout the book, we have been reminded of the importance of not drawing causal conclusions from bivariate relationships, since it is always possible that important variables have been excluded from the analysis. Multiple regression controls for several IVs and thus helps minimize the threat that omitted variables pose a threat to the validity of the findings. Multiple regression isolates the effect of each IV so that the effect of a one-unit increase of an individual IV on the DV can be determined, while holding all other variables constant.

5.

a. $b = \dfrac{5(86)-24(19)}{5(230)-24^2} = \dfrac{430-456}{1150-576} = -.05$

5b. a = 3.80 − (-.05)4.80 = 3.80 − (-.24) = 4.04

5c. $\hat{y} = 4.04 − .05x$

5d. i. For x = 3, $\hat{y} = 4.04 − .05(3) = 3.89$
ii. For x = 15, $\hat{y} = 4.04 − .05(15) = 3.26$

5e. Step 1: H_0: B = 0 and H_1: B ≠ 0
Step 2: t distribution with df = 5 − 2 = 3
Step 3: t_{crit} = ±3.182 and the decision rule is: If t_{obt} is greater than 3.182 or less than -3.182, the null will be rejected.