A brand manager in a large packaged goods company is interested in understanding the impact of price and promotional activity on the relative share of his product (which we will call here "Brand X"). The file *BRAND_X* contains 101 weeks of data from a panel of consumers who kept track of their shopping trips and purchase activities in the product category. The variables in the data set are defined in Table 3.16.

a. Construct a linear regression model of the market share of Brand X. Justify your choice of variables as well as any transformations you choose to use.
b. Estimate the parameters of your model. How well does the model fit? Does it explain a significant amount of the observed variation in Brand X's market share?
c. All other things remaining equal, is the share of Brand X increasing, decreasing, or holding steady over time?
d. How would you describe the effect of price on the share of Brand X?
e. In week 157, Brand X achieves a relative price index of 100 and a promotion index of 30. The number of shopping trips by loyal households is 35; the number of shopping trips by non-loyal households is 60. What is your forecast of the market share for Brand X during that week?

TABLE 3.16 Variables in *BRAND_X* data set

Variable	Definition
SHARE_X	Market share for Brand X (in percent)
WEEK	Week (from 56 to 156)
REL_PRICE	Relative price index for Brand X (defined as the price of Brand X divided by the average price of all competing brands × 100)
LAG_PRICE	Lagged price index for Brand X
PROMO	Index of promotional activity (i.e., special deals other than price) by Brand X: 0 = no promotional activity by Brand X; 100 = all promotional activity is Brand X
LOY_HH	Number of shopping trips taken by households in the panel characterized as "loyal" to Brand X
NONLOY_HH	Number of shopping trips taken by households in the panel characterized as "not loyal" to Brand X

The data in the file *BANK_WAGES* are from a study of employee compensation at the Harris Bank (Roberts, 1979). The file contains information on 116 employees (35 white males and 81 white females) hired during the years 1969–71 as general office trainees. The variables in the data set are defined in Table 3.17.

a. How could you use regression to test for the presence of gender discrimination in compensation by the bank? What should the regression model look like?
b. Test the hypothesis that there is no difference between men and women in the level of compensation offered by the bank. Be sure to control for differences in education, prior work experience, and seniority. What is your conclusion?
c. Examine the residuals from your regression model. Are there any influential observations in the data? What other concerns might you have about the model?

TABLE 3.17 Variables in *BANK_WAGES* data set

Variable	Definition
SAL_NOW	Salary (in dollars) at the time of the study (March, 1977)
ID	Employee ID
SAL_START	Salary (in dollars) at the time of hiring
SEX	Sex of employee (0 = male, 1 = female)
SENIOR	Seniority (in months)
AGE	Age (in months) at time of hire
WORK_EXP	Months of work experience
EDUC	Years of education
HS_GRAD	Indicator (1 = high school graduate)
COLLEGE	Indicator (1 = some college)
COLL_GRAD	Indicator (1 = college graduate)
GRAD_SCH	Indicator (1 = any graduate work)

Ofir and Simonson (2001) conducted several experiments investigating the effect of "expecting to evaluate" on ratings of product quality. They found that in a variety of situations, consumers are more critical (and provide more negative ratings) when they are told in advance that they will be asked to provide an evaluation after they have experienced a product or service.

We now consider a subset of the data collected by Ofir and Simonson in one of their studies. A group of 201 subjects was asked to read an article that "appeared in the first issue of a new magazine," and all subjects were told that "we will ask for your evaluation of the writing quality of the magazine based on this article." Thus, all subjects expect to evaluate the article after reading it. For this group, there were two experimental treatments: the quality of the article (high, low) and expectations about the quality of the article (high, low). To accomplish the first manipulation, Ofir and Simonson took an article from *The New York Times*. The "high-quality" treatment was the original article; the "low-quality" version was created by using slang and poor grammar without changing the content. They manipulated expectations of quality by telling subjects the new magazine was either "started by experienced journalists" (high expectations) or "started by freshman students at a local high school" (low expectations). For the dependent variable, Ofir and Simonson asked, "What is your evaluation of the magazine based on the professionalism and quality of the writing, grammar, language, and editing of the article?" Subjects responded on a seven-point scale from "Very favorable" to "Very unfavorable."

Data from the experiment are available in the file *EXPECT_EVAL*. There are six columns of data in the file, described below:

Col 1	Subject ID
Col 2	Evaluation manipulation (a = expecting to evaluate)
Col 3	Quality manipulation (b = bad quality, g = good quality)
Col 4	Expectations manip (h = expect high quality, l = expect low quality)
Col 5	Y_1 (evaluation of magazine, 7-point scale)
Col 6	Y_2 (agreement with issue in the article, 7-point scale)

a. Using ANOVA, test the null hypothesis that there are no differences in the evaluation of the magazine across experimental treatment groups. Can you reject the hypothesis at the 0.05 level? If so, which of the treatment effects is significant?

b. Describe the nature of the interaction between quality and expected quality.

Ofir and Simonson collected additional information from subjects regarding their "need for cognition." They collected a set of 18 pretested measures all designed to measure the subject's need for cognition (the variables X_3, X_4, X_5, X_7, X_8, X_9, X_{12}, X_{16}, and X_{17} were reverse-coded). These data are available in the file *COGNITION* (and described in more detail in Chapter 4).

c. Using the measures in the file *COGNITION*, create a "need for cognition" index denoted Z. Introduce Z into the ANOVA model in part a as a covariate, and retest the hypothesis from part a above. How do your results change (if at all)?

Ofir and Simonson collected a second evaluation measure that was designed to capture the persuasiveness of the editorial message (i.e., issue agreement). We denote this second evaluation measure as Y_2.