

Problems for student projects

8–5. A real estate development firm wants to test whether there are differences in the average price of a lot of a given size in the center of each of four cities: Philadelphia, New York, Washington, and Baltimore. Random samples of 52 lots in Philadelphia, 38 lots in New York, 43 lots in Washington, and 47 lots in Baltimore lead to a computed test statistic value of 12.53. Do you believe that average lot prices among the four cities are equal? How confident are you of your conclusion? Explain.

8–19. In the theory of finance, a market for any asset or commodity is said to be *efficient* if items of identical quality and other attributes (such as risk, in the case of stocks) are sold at the same price. A Geneva-based oil industry analyst wants to test the hypothesis that the spot market for crude oil is efficient. The analyst chooses the Rotterdam oil market, and he selects Arabian Light as the type of oil to be studied. (Differences in location may cause price differences because of transportation costs, and differences in the type of oil—hence, in the quality of oil—also affect the price. Therefore, both the type and the location must be fixed.) The analyst also notes that during the month of February 1987, the “official” price of a barrel of crude oil did not change. A random sample of eight observations from each of four sources of the spot price of a barrel of oil during February is collected. Data, in U.S. dollars per barrel, are as follows.

U.K.	Mexico	U.A.E.	Oman
\$17.80	\$18.01	\$18.10	\$18.05
18.00	17.75	17.92	18.01
17.98	18.00	18.01	17.94
18.20	17.77	17.88	18.23
18.00	18.01	18.30	18.20
17.99	18.01	18.22	18.00
18.10	18.12	18.56	17.84
17.90	18.20	18.10	18.11

Based on these data, what should the analyst conclude about whether or not the market for crude oil is efficient? Are conclusions valid only for the Rotterdam market? Are conclusions valid for all types of crude oil? What assumptions are necessary for the analysis? Do you believe that all the assumptions are met in this case? What are the limitations, if any, of this study? Discuss.

8–21. An article in the *Journal of Marketing* describes a study of organizational buyer behavior.⁸ As part of the analysis, the authors of the study wanted to test whether or not differences exist in the average amount of time buyers need for making a group-purchase decision about one of three different kinds of copying machines. The article reports that the degrees of freedom for error were 82, the mean square treatment was 136.3, and the sum of squares total was 1,701.83. Construct the complete ANOVA table for this study, and determine whether or not there is evidence of differences in the average time to reach a group-purchase decision about copying machines.

- 8–23. A study is undertaken to determine whether differences exist in average consumer quality ratings of the following brands of color television sets: Magnavox, General Electric, Panasonic, Zenith, Sears, Philco, Sylvania, and RCA. For each brand, 100 randomly chosen consumer responses are available, and from these the following quantities are computed: $SSTR = 45,210$; $SSTO = 92,340$. Construct an ANOVA table for this study, and test the null hypothesis that all eight brands have equal average consumer quality ratings versus the alternative hypothesis that they do not.
- 9–13. A financial analyst at Goldman Sachs ran a regression analysis of monthly returns on a certain investment (Y) versus returns for the same month on the Standard & Poor's Index (X). The regression results included $SS_X = 765.98$ and $SS_{XY} = 934.49$. Give the least-squares estimate of the regression slope parameter.
- 9–29. Is it possible that a sample correlation of 0.51 between two variables will not indicate that the two variables are really correlated, while a sample correlation of 0.04 between another pair of variables will be statistically significant? Explain.
- 9–33. A study was conducted to determine whether a correlation exists between consumers' perceptions of a television commercial (measured on a special scale) and their interest in purchasing the product (measured on a scale). The results are $n = 65$, $r = 0.37$. Is there statistical evidence of a correlation between the two variables?
- 9–35. A regression analysis of fuel efficiency (X) versus sales (Y) of different types of corporate aircraft includes the following results: $b_1 = 2.435$, $s(b_1) = 1.567$, $n = 12$. Do you believe there is a linear relationship between sales of corporate aircraft and the aircraft's fuel efficiency?
- 9–41. An advertising research firm conducts a study to determine the relationship between the length of a commercial and the resulting product recall scores of people viewing the commercial. Results of the analysis include $b_1 = 3.467$, $s(b_1) = 0.775$, and $n = 23$. Is there evidence of a linear relationship between commercial length and viewer recall score?
- 9–43. A management recruiter wants to estimate a linear regression relationship between an executive's experience and the salary the executive may expect to earn after placement with an employer. From data on 28 executives, which are assumed a random sample from the population of executives that the recruiter places, the following regression results are obtained: $b_1 = 5.49$, $s(b_1) = 1.21$. Is there a linear relationship between the experience and the salary of executives placed by the recruiter?
- 9–59. In a simple linear regression analysis, it is found that $b_1 = 2.556$ and $s(b_1) = 4.122$. The sample size is $n = 22$. Conduct an F test for the existence of a linear relationship between the two variables.
- 9–75. The computer output for a simple linear regression of sales of one product versus sales of another product includes the following: $n = 128$, $SSE = 3,343$, $SSR = 45,678$. Compute the remaining sum of squares, the mean squares, the F ratio, and the coefficient of determination. Write an ANOVA table for this regression. What can you say about the relationship between sales of the two products?

- 9–83. A car rental agency is interested in studying the relationship between the mileage of their subcompact cars (X) and the monthly maintenance cost for the cars (Y). Data on 15 randomly chosen cars are given below. Carry out a simple linear regression analysis of the relationship between the two variables. What is the expected monthly maintenance cost for a car with 13,000 miles?

Miles (in thousands):	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Cost (dollars):	13	16	15	20	19	21	26	24	30	32	30	35	34	40	39

- 9–77. A regression analysis of electricity use (Y) versus temperature (X) in a city in the Northwest includes the following results: $SSE = 5,432.87$, $SST = 87,695.98$; the sample size is 23 (days). Conduct the F test for the existence of a linear relationship between the two variables. Do you believe temperature has an effect on the consumption of electricity?
- 9–79. The amount of money a manufacturer is willing to pay for shipping its goods is believed to be inversely related to the length of time the goods are in transit. A shipping company analyst carries out a simple linear regression analysis of the time in transit (X) versus shipping costs (Y) and obtains: $b_1 = -3.453$, $s(b_1) = 0.987$. Conduct the t test for the existence of a linear relationship between shipping costs and length of transit time. From this information, can you also give the F statistic value? (The sample size is $n = 16$ transatlantic crossings.)
- 9–81. A computer software manufacturer wants to study the relationship between the number of microcomputers in use in different areas and the number of software packages the company sells in the areas. A simple linear regression analysis of 21 geographical regions reveals the following: $b_0 = 12.43$, $b_1 = 1.076$, $s(b_0) = 13.65$, $s(b_1) = 0.083$, $SSE = 1,076.11$, $SS_x = 72.641$, $\bar{x} = 12,453$ computers. Using this information, construct 95% confidence intervals for the regression parameters β_0 and β_1 , a 95% confidence interval for the *average* number of software packages sold in an area where 10,000 computers are in use, and a 95% prediction interval for the number of packages sold in that area.
- 10–13. Avis is interested in estimating weekly costs of maintenance of its rental cars of a certain size based on the variables: number of miles driven during the week, number of renters during the week, the car's total mileage, and the car's age. A regression analysis is carried out, and the results include: $n = 45$ cars (each car selected randomly, during a randomly selected week of operation), $SSR = 7,768$, and $SSTO = 15,673$. Construct a complete ANOVA table for this problem, and test for the existence of a linear regression relationship between weekly maintenance costs and any of the four independent variables considered.
- 10–23. In the Nissan Motor Company situation in problem 10–21, suppose that a new variable is considered for inclusion in the equation and a new regression relationship is analyzed with the new variable included. Suppose that the resulting multiple coefficient of determination is $R^2 = 91.8\%$. Find the adjusted multiple coefficient of determination. Should the new variable be included in the final regression equation? Give your reasons for including or excluding the variable.
- 10–27. In a regression analysis with six independent variables and a data set of 250 points, it is found that $SSE = 5,445$ and $SSTO = 22,679$. Construct an ANOVA table, conduct the F test, find R^2 and $\bar{R}^2$, and find MSE.