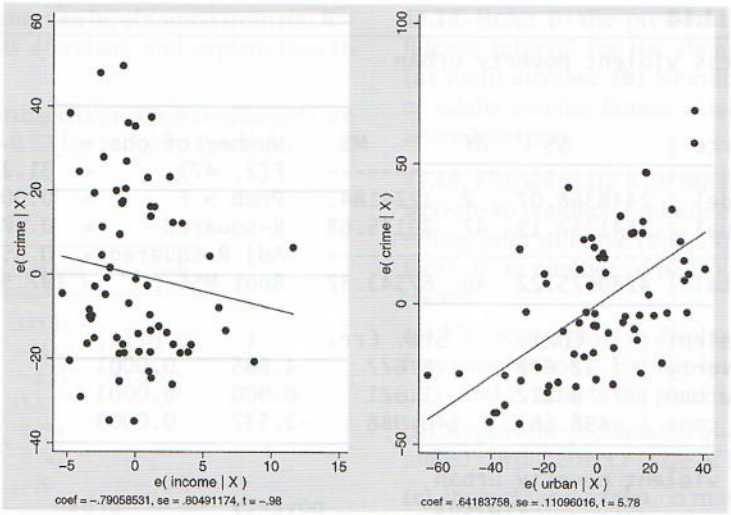

FIGURE 11.13

TABLE 11.13

BIVARIATE	Coef.	Std. Err.	t	P>|t|
income	2.609	0.675	3.866	0.0003
_cons	-11.526	16.834	-0.685	0.4960
MULTIPLE	**Coef.**	**Std. Err.**	**t**	**P>|t|**
income	-0.809	0.805	-1.005	0.3189
urban	0.646	0.111	5.811	0.0001
_cons	40.261	16.365	2.460	0.0166

(c) Find R^2 for the multiple regression model, and show that it is not much larger than r^2 for the model using urbanization alone as the predictor. Interpret.

11.9. Recent UN data from several nations on y = crude birth rate (number of births per 1000 population size), x_1 = women's economic activity (female labor force as percentage of male), and x_2 = GNP (per capita, in thousands of dollars) has prediction equation $\hat{y} = 34.53 - 0.13x_1 - 0.64x_2$. The bivariate prediction equation with x_1 is $\hat{y} = 37.65 - 0.31x_1$. The correlations are $r_{yx_1} = -0.58$, $r_{yx_2} = -0.72$, and $r_{x_1x_2} = 0.58$. Explain why the coefficient of x_1 in the bivariate equation is quite different from that in the multiple predictor equation.

11.10. For recent UN data for several nations, a regression of carbon dioxide use (CO_2, a measure of air pollution) on gross domestic product (GDP) has a correlation of 0.786. With life expectancy as a second explanatory variable, the multiple correlation is 0.787.

(a) Explain how to interpret the multiple correlation.

(b) For predicting CO_2, did it help much to add life expectancy to the model? Does this mean that life expectancy is very weakly correlated with CO_2? Explain.

11.11. Table 11.14 shows Stata output from fitting the multiple regression model to recent statewide data, excluding D.C., on y = violent crime rate (per 100,000 people), x_1 = poverty rate (percentage with income below the poverty level), and x_2 = percentage living in urban areas.

(a) Report the prediction equation.

(b) Massachusetts had $y = 805$, $x_1 = 10.7$, and $x_2 = 96.2$. Find its predicted violent crime rate. Find the residual, and interpret.

(c) Interpret the fit by showing the prediction equation relating $\hat{y}$ and x_1 for states with (i) $x_2 = 0$, (ii) $x_2 = 100$. Interpret.

(d) Interpret the correlation matrix.

(e) Report R^2 and the multiple correlation, and interpret.

11.12. Refer to the previous exercise.

(a) Report the F statistic for testing H_0: $\beta_1 = \beta_2 = 0$, report its df values and P-value, and interpret.

(b) Show how to construct the t statistic for testing H_0: $\beta_1 = 0$, report its df and P-value for H_a: $\beta_1 \neq 0$, and interpret.

(d) Explain how the F-value can be obtained from the R^2-value reported. Report its df values, and explain how to interpret its result.

(e) The estimated standardized regression coefficients are -0.79 for ideology and -0.23 for religion. Interpret.

TABLE 11.17

Variable	Coefficient	Std. Error
Intercept	135.31	
Ideology	−14.07	3.16**
Religion	−2.95	2.26
F	13.93**	
R^2	0.799	
Adj. R^2	0.742	
n	10	

11.16. Refer to Table 11.5 on page 328. Test H_0: $\beta_2 = 0$ that mental impairment is independent of SES, controlling for life events. Report the test statistic, and report and interpret the P-value for **(a)** H_a: $\beta_2 \neq 0$, **(b)** H_a: $\beta_2 < 0$.

11.17. For a random sample of 66 state precincts, data are available on y = percentage of adult residents who are registered to vote, x_1 = percentage of adult residents owning homes, x_2 = percentage of adult residents who are nonwhite, x_3 = median family income (thousands of dollars), x_4 = median age of residents, x_5 = percentage of residents who have lived in the precinct for at least 10 years. Table 11.18 shows some output used to analyze the data.

(a) Fill in all the missing values.

(b) Do you think it is necessary to include all five explanatory variables in the model? Explain.

(c) To what test does "F value" refer? Interpret the result of that test.

(d) To what test does the t-value opposite x1 refer? Interpret the result of that test.

TABLE 11.18

	Sum of Squares	DF	Mean Square	F	Sig
Regression	----	---	----	----	----
Residual	2940.0	---	----		
Total	3753.3	---			

R-Square ----

Root MSE ----

Variable	Parameter Estimate	Standard Error	t	Sig
Intercept	70.0000			
x1	0.1000	0.0450	----	----
x2	-0.1500	0.0750	----	----
x3	0.1000	0.2000	----	----
x4	-0.0400	0.0500	----	----
x5	0.1200	0.0500	----	----

11.18. Refer to the previous exercise. Find a 95% confidence interval for the change in the mean of y for a **(a)** 1-unit increase, **(b)** 50-unit increase in the percentage of adults owning homes, controlling for the other variables. Interpret.

11.19. Use software with the Houses data file at the text website to conduct a multiple regression analysis of y = selling price of home (dollars), x_1 = size of home (square feet), x_2 = number of bedrooms, x_3 = number of bathrooms.

(a) Use scatterplots to display the effects of the explanatory variables on y. Explain how the highly discrete nature of x_2 and x_3 affects the plots.

(b) Report the prediction equation and interpret the estimated partial effect of size of home.

(c) Inspect the correlation matrix, and report the variable having the (i) strongest association with y, (ii) weakest association with y.

(d) Report R^2 for this model and r^2 for the simpler model using x_1 alone as the explanatory variable. Interpret.

11.20. Refer to the previous exercise.

(a) Test the partial effect of number of bathrooms, and interpret.

(b) Find the partial correlation between selling price and number of bathrooms, controlling for number of bedrooms. Compare it to the correlation, and interpret.

(c) Find the estimated standardized regression coefficients for the model, and interpret.

(d) Write the prediction equation using standardized variables. Interpret.

11.21. Exercise 11.11 showed a regression analysis for statewide data on y = violent crime rate, x_1 = poverty rate, and x_2 = percentage living in urban areas. When we add an interaction term, we get $\hat{y} = 158.9 - 14.72x_1 - 1.29x_2 + 0.76x_1x_2$.

(a) As the percentage living in urban areas increases, does the effect of poverty rate tend to increase or decrease? Explain.

(b) Show how to interpret the prediction equation, by finding how it simplifies when $x_2 = 0$, 50, and 100.

11.22. A study analyzes relationships among y = percentage vote for Democratic candidate, x_1 = percentage of registered voters who are Democrats, and x_2 = percentage of registered voters who vote in the election, for several congressional elections in 2016. The researchers expect interaction, since they expect a higher slope between y and x_1 at larger values of x_2 than at smaller values. They obtain the prediction equation $\hat{y} = 20 + 0.30x_1 + 0.05x_2 + 0.005x_1x_2$. Does this equation support the direction of their prediction? Explain.

11.23. Use software with the Houses data file to allow interaction between number of bedrooms and number of bathrooms in their effects on selling price.

(a) Interpret the fit by showing the prediction equation relating $\hat{y}$ and number of bedrooms for homes with (i) two bathrooms, (ii) three bathrooms.

(b) Test the significance of the interaction term. Interpret.

11.24. A multiple regression analysis investigates the relationship between y = college GPA and several explanatory variables, using a random sample of 195 students at Slippery Rock University. First, high school GPA and total SAT score are entered into the model. The sum of squared errors is SSE = 20. Next, parents' education and parents' income are added, to determine if they have an effect, controlling for high school GPA and SAT. For this expanded model, SSE = 19. Test whether this complete model is significantly better than the one containing only high school GPA and SAT. Report and interpret the P-value.

11.25. Table 11.19 shows results of regressing y = birth rate (number of births per 1000 population) on x_1 = women's economic activity and x_2 = literacy rate, using UN data for 23 nations.

(a) Report the value of each of the following:
(i) r_{yx_1}, (ii) r_{yx_2}, (iii) R^2,
(iv) TSS, (v) SSE, (vi) mean square error,
(vii) s, (viii) s_y, (ix) se for b_1,
(x) t for H_0: $\beta_1 = 0$,
(xi) P for H_0: $\beta_1 = 0$ against H_a: $\beta_1 \neq 0$,
(xii) P for H_0: $\beta_1 = 0$ against H_a: $\beta_1 < 0$,
(xiii) F for H_0: $\beta_1 = \beta_2 = 0$,
(xiv) P for H_0: $\beta_1 = \beta_2 = 0$.

(b) Report the prediction equation, and interpret the signs of the estimated regression coefficients.

(c) Interpret the correlations r_{yx_1} and r_{yx_2}.

(d) Report R^2, and interpret its value.

TABLE 11.19

	Mean	Std Deviation	N
BIRTHS	22.117	10.469	23
ECON	47.826	19.872	23
LITERACY	77.696	17.665	23

Correlations

Correlation		BIRTHS	ECON	LITER
	BIRTHS	1.00000	-0.61181	-0.81872
	ECON	-0.61181	1.00000	0.42056
	LITERACY	-0.81872	0.42056	1.00000

	Sum of Squares	DF	Mean Square	F	Sig
Regression	1825.969	2	912.985	31.191	0.0001
Residual	585.424	20	29.271		
Total	2411.393	22			

Root MSE (Std. Error of the Estimate) 5.410 R Square 0.7572

	Unstandardized Coeff. B	Std. Error	Standardized Coeff. (Beta)	t	Sig
(Constant)	61.713	5.2453		11.765	0.0001
ECON	-0.171	0.0640	-0.325	-2.676	0.0145
LITERACY	-0.404	0.0720	-0.682	-5.616	0.0001