4.34 Assertiveness and leaderships. Management professors at Columbia University examined the relationship between assertiveness and leadership (*Journal of Personality and Social Psychology*, February 2007). The sample was comprised of 388 people enrolled in a full-time MBA program. Based on answers to a questionnaire, the researchers measured two variables for each subject: assertiveness score (x) and leadership ability score (y). A quadratic regression model was fit to the data with the following results:

INDEPENDENT VARIABLE	β ESTIMATE	t-VALUE	p-VALUE
x	.57	2.55	.01
x^2	−.88	−3.97	< .01
Model $R^2 = .12$			

(a) Conduct a test of overall model utility. Use $\alpha = .05$.

$H_0: \beta_1 = \beta_2 = 0, H_a$: at least one of the above coefficient is nonzero.

$$F_{\alpha,k,n-(k+1)} = F_{0.05,2,388-(2+1)} = F_{0.05,2,385} \approx 3.019$$

(Looked up from https://www.danielsoper.com/statcalc/calculator.aspx?id=4)

$$F_{stat} = \frac{R^2/k}{(1-R^2)/(n-(k+1))} = \frac{.12/2}{(1-.12)/385} = \frac{.06}{.00229} = 26.25$$

Because $F_{stat} > F_{0.05,2,385}$, so we reject H_0, and at least one of the above is nonzero.

Ans: At least one of $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ is nonzero

(b) The researchers hypothesized that a leadership ability will increase at a decreasing rate with assertiveness, set up the null and alternative hypothesis to test this theory.

$H_0: \beta_2 = 0, H_a: \beta_2 < 0$

Ans: $\boldsymbol{H_0: \beta_2 = 0, H_a: \beta_2 < 0}$

(c) Use the reported results to conduct the test, part (b). Give your conclusion (at $\alpha = .05$) in the words of the problem.

$t_{\alpha,n-(k+1)} = t_{0.05,388-(2+1)} = t_{0.05,385} = -1.6488$ (one tail)

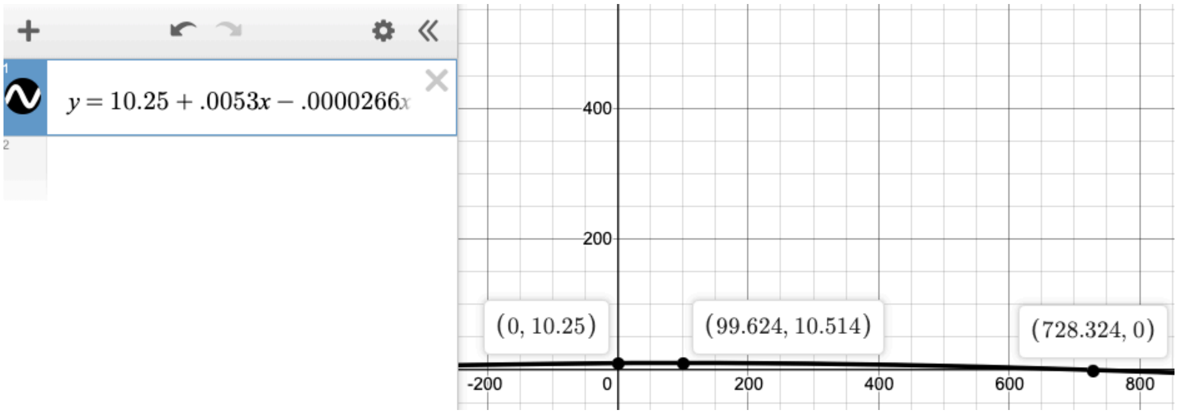
(Looked up from http://www.ttable.org/student-t-value-calculator.html)

$t_{stat} = -3.97 < -1.6488$, so we reject H_0, and there is very strong evidence of downward curvature in the population.

Ans: Leadership will increase at a decreasing rate with assertiveness

4.38 Estimating change-point dosage. A standard method for studying toxic substances and their effects on humans is to observe the responses of rodents exposed to various doses of the substance over time. In the *Journal of Agricultural, Biological, and Environmental Statistics* (June 2005), researchers used least squares regression to estimate the "change-point" dosage—defined as the largest dose level that has no adverse effects. Data were obtained from a dose–response study of rats exposed to the toxic substance aconiazide. A sample of 50 rats was evenly divided into five dosage groups: 0, 100, 200, 500, and 750 milligrams per kilograms of body weight. The dependent variable y measured was the weight change (in grams) after a 2-week exposure. The researchers fit the quadratic model $E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$, where x = dosage level, with the following results: $\hat{y} = 10.25 + .0053x - .0000266x^2$.

(a) Construct a rough sketch of the least squares prediction equation. Describe the nature of the curvature in the estimated model.



Observation: a downward curvature

Ans: downward curvature

(b) Estimate the weight change (y) for a rat given a dosage of 500 mg/kg of aconiazide.

$$\hat{y} = 10.25 + .0053(500) - .0000266(500^2) = 6.25$$

Ans: 6.25 g

(c) Estimate the weight change (y) for a rat given a dosage of 0 mg/kg of aconiazide. (This dosage is called the "control" dosage level.)

$$\hat{y} = 10.25 + .0053(0) - .0000266(0^2) = 10.25$$

Ans: 10.25 g

(d) Of the five groups in the study, find the largest dosage level x that yields an estimated weight change that is closest to but below the estimated weight change for the control group. This value is the "change-point" dosage.

$$\hat{y} = 10.25 + .0053(0) - .0000266(0^2) = 10.25$$
$$\hat{y} = 10.25 + .0053(100) - .0000266(100^2) = 10.514$$
$$\hat{y} = 10.25 + .0053(200) - .0000266(200^2) = 10.246$$
$$\hat{y} = 10.25 + .0053(500) - .0000266(500^2) = 6.25$$
$$\hat{y} = 10.25 + .0053(750) - .0000266(750^2) = -0.738$$

Ans: 200 mg/kg

4.64 Cooling method for gas turbines. Refer the *Journal of Engineering for Gas Turbines and Power* (January 2005) study of a high-pressure inlet fogging method for a gas turbine engine, Exercise 4.13 (p. 188). Consider a model for heat rate (kilojoules per kilowatt per hour) of a gas turbine as a function of cycle speed (revolutions per minute) and cycle pressure ratio. The data are saved in the GASTURBINE file.

(a) Write a complete second-order model for heat rate (y).

Ans: $\boldsymbol{\hat{y} = 15585 + 0.078x_1 - 523x_2 + 0.00445x_1x_2 - 0.000000x_1^2 + 8.84x_2^2}$

where x_1 = cycle speed (revolutions per minute) and x_2 = cycle pressure ratio.

(b) Give the null and alternative hypotheses for determining whether the curvature terms in the complete second-order model are statistically useful for predicting heat rate (y).

Ans: $\boldsymbol{H_0: \beta_4 = \beta_5 = 0, H_a}$**: Either** $\boldsymbol{\beta_4}$ **or** $\boldsymbol{\beta_5}$ **(or both) are nonzero.**

(c) For the test in part (b), identify the "complete" and "reduced" model.

Complete model: $\boldsymbol{\hat{y} = \beta_0 + \beta_1x_1 - \beta_2x_2 + \beta_3x_1x_2 + \beta_4x_1^2 + \beta_5x_2^2}$

Reduced model: $\boldsymbol{\hat{y} = \beta_0 + \beta_1x_1 - \beta_2x_2 + \beta_3x_1x_2}$

(d) Portions of the MINITAB printouts for the two models are shown below. Find the values of SSE_R, SSE_C, and MSE_C on the printouts.

Ans: SSE_R = 25310639, SSE_C = 19370350, MSE_C = 317547

MINITAB Output for Exercise 4.64

Complete Model

```
The regression equation is
HEATRATE = 15583 + 0.078 RPM - 523 CPRATIO + 0.00445 RPM_CPR - 0.000000 RPMSQ
           + 8.84 CPRSQ

S = 563.513   R-Sq = 88.5%   R-Sq(adj) = 87.5%

Analysis of Variance

Source          DF         SS        MS      F      P
Regression       5  148526859  29705372  93.55  0.000
Residual Error  61   19370350    317547
Total           66  167897208
```

Reduced Model

```
The regression equation is
HEATRATE = 12065 + 0.170 RPM - 146 CPRATIO - 0.00242 RPM_CPR

S = 633.842   R-Sq = 84.9%   R-Sq(adj) = 84.2%

Analysis of Variance

Source          DF         SS        MS       F      P
Regression       3  142586570  47528857  118.30  0.000
Residual Error  63   25310639    401756
Total           66  167897208
```

(e) Compute the value of the test statistics for the test of part (b).

$$F = \frac{(\mathrm{SSE_R} - \mathrm{SSE_C})/2}{\mathrm{MSE_C}} = \frac{(25310639 - 19370350)/2}{317547} = \frac{2970144.5}{317547} = 9.353$$

Ans: *F* = 9.353

(f) Find the rejection region for the test of part (b) using $\alpha = .10$.

$n - (k+1) = 66 - (5+1) = 60, F_{0.10,2,60} = 2.39$ (from Table D.3)

Ans: 2.39

(g) State the conclusion of the test in the words of the problem.

Ans: The quadratic terms contribute to the prediction of *y*, the heat rate.

4.68 Buy-side versus sell-side analysts' earnings forecasts. Refer to the *Financial Analysts Journal* (July/August 2008) comparison of earnings forecasts of buy-side and sell-side analysts, Exercise 4.56 (p. 224). Recall that the Harvard Business School professors used regression to model the relative optimism (y) of the analysts' 3-month horizon forecasts as a function of x_1 = {1 if the analyst worked for a buy-side firm, 0 if the analyst worked for a sell-side firm} and x_2 =

number of days between forecast and fiscal year-end (i.e., forecast horizon). Consider the complete second-order model

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 (x_2)^2 + \beta_5 x_1 (x_2)^2$$

(a) What null hypothesis would you test to determine whether the quadratic terms in the model are statistically useful for predicting relative optimism (y)?

Ans: $H_0: \beta_4 = \beta_5 = 0, H_a$: Either β_4 or β_5 (or both) are nonzero.

(b) Give the complete and reduced models for conducting the test, part (a).

Complete model: $\hat{y} = \beta_0 + \beta_1 x_1 - \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2$

Reduced model: $\hat{y} = \beta_0 + \beta_1 x_1 - \beta_2 x_2 + \beta_3 x_1 x_2$

(c) What null hypothesis would you test to determine whether the interaction terms in the model are statistically useful for predicting relative optimism (y)?

Ans: $H_0: \beta_3 = 0; H_a: \beta_3 > 0.$

(d) Give the complete and reduced models for conducting the test, part (c).

Complete model: $\hat{y} = \beta_0 + \beta_1 x_1 - \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2$

Reduced model: $\hat{y} = \beta_0 + \beta_1 x_1 - \beta_2 x_2$

(e) What null hypothesis would you test to determine whether the dummy variable terms in the model are statistically useful for predicting relative optimism (y)?

Ans: $H_0: \beta_1 = 0; H_a: \beta_1 \neq 0.$

(f) Give the complete and reduced models for conducting the test, part (e).

Complete model: $\hat{y} = \beta_0 + \beta_1 x_1 - \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2$

Reduced model: $\hat{y} = \beta_0 + \beta_1 x_1$

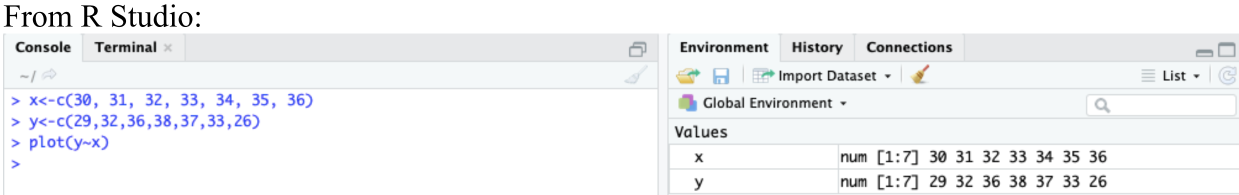
5.10 Tire wear and pressure. Underinflated or overinflated tires can increase tire wear and decrease gas mileage. A new tire was tested for wear at different pressures with the results shown in the table.

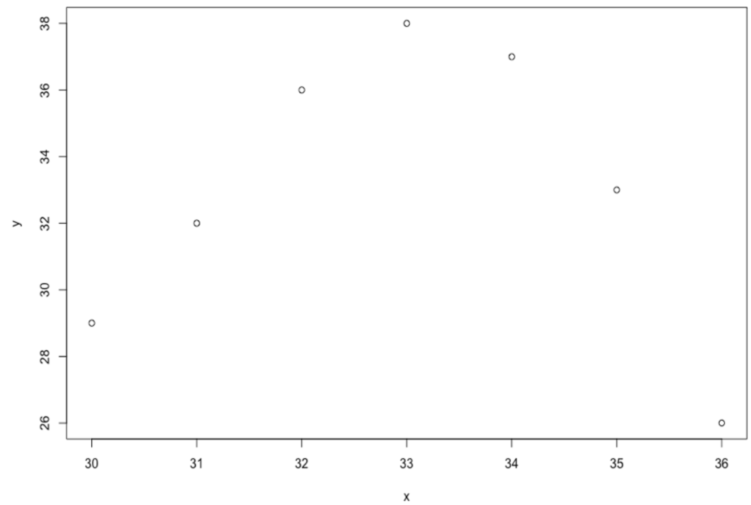
TIRES2

PRESSURE	MILEAGE
x, pounds per square inch	y, thousands
30	29
31	32
32	36
33	38
34	37
35	33
36	26

(a) Graph the data in a scatterplot.

From R Studio:



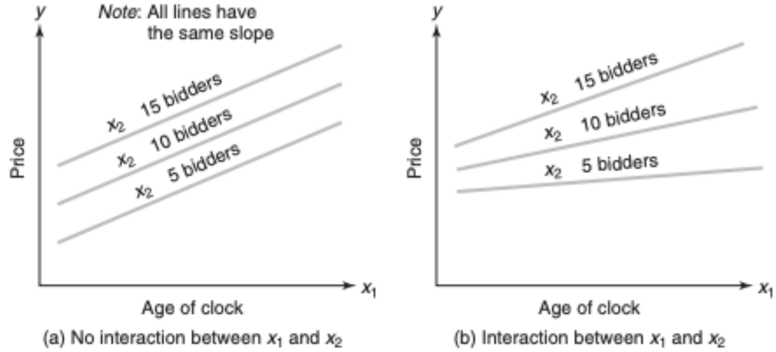


(b) If you were given the information for x = 30, 31, 32, and 33 only, what kind of model would you suggest? For x = 33, 34, 35, and 36? For all the data?
***x* = 30, 31, 32, and 33 only: first-order (straight-line) model with a positive slope**
***x* = 33, 34, 35, and 36: first-order (straight-line) model with a negative slope**
all data: second-order (quadratic) model

5.12 Signal-to-noise ratios of seismic waves. Chinese scientists have developed a method of boosting the signal-to-noise ratio of a seismic wave (*Chinese Journal of Geophysics,* Vol. 49, 2006). Suppose an exploration seismologist wants to develop a model to estimate the average signal-to-noise ratio of an earthquake's seismic wave, y, as a function of two independent variables:
x_1 = Frequency (cycles per second), x_2 = Amplitude of the wavelet

(a) Identify the independent variables as quantitative or qualitative.
Ans: Both x_1 and x_2 are quantitative.
(b) Write the first-order model for $E(y)$.
Ans: $\hat{\boldsymbol{y}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_1} + \boldsymbol{\beta_2 x_2}$
(c) Write a model for $E(y)$ that contains all first-order and interaction terms. Sketch typical response curves showing $E(y)$, the mean signal-to-noise ratio, versus x_2, the amplitude of the wavelet, for different values of x_1 (assume that x_1 and x_2 interact).
Sketch will be similar to Figure 4.10 from textbook with different slopes.
Ans: $\hat{\boldsymbol{y}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_1} + \boldsymbol{\beta_2 x_2} + \boldsymbol{\beta_3 x_1 x_2}$

Figure 4.10 Examples of no-interaction and interaction models



(d) Write the complete second-order model for $E(y)$.

Ans: $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_2^2$

5.13 Signal-to-noise ratios of seismic waves (Cont'd). Refer to Exercise 5.12. Suppose the model from part (c) is fit, with the following result: $\hat{y} = 1 + .05x_1 + x_2 + .05x_1x_2$
Graph the estimated signal-to-noise ratio $\hat{y}$ as a function of the wavelet amplitude, x_2, over the range $x_2 = 10$ to $x_2 = 50$ for frequencies of $x_1 = 1$, 5, and 10. Do these functions agree (approximately) with the graphs you drew for Exercise 5.12, part (c)?
Ans: $x_1 = 1$: bottom; $x_1 = 5$: middle; $x_1 = 10$: top. These functions with the graph in Exercise 5.12(c).