Chapter 12 Selected Problem Solutions

Section 12-1

12-1. a) $$X'X = \begin{bmatrix} 10 & 223 & 553 \\ 223 & 5200.9 & 12352 \\ 553 & 12352 & 31729 \end{bmatrix}$$

$$X'y = \begin{bmatrix} 1916.0 \\ 43550.8 \\ 104736.8 \end{bmatrix}$$

b) $\hat{\beta} = \begin{bmatrix} 171.054 \\ 3.713 \\ -1.126 \end{bmatrix}$, so $\hat{y} = 171.054 + 3.714x_1 - 1.126x_2$

c) $\hat{y} = 171.054 + 3.714(18) - 1.126(43) = 189.481$

12-5.

```
Predictor       Coef       StDev          T        P
Constant      33.449       1.576      21.22    0.000
x1         -0.054349    0.006329      -8.59    0.000
x6            1.0782      0.6997       1.54    0.138
S = 2.834       R-Sq = 82.9%     R-Sq(adj) = 81.3%
Analysis of Variance
Source        DF         SS          MS         F        P
Regression     2     856.24      428.12     53.32    0.000
Error         22     176.66        8.03
Total         24    1032.90
```

a) $\hat{y} = 33.4491 - 0.05435x_1 + 1.07822x_2$

b) $\hat{\sigma}^2 = 8.03$

c) $\hat{y} = 33.4491 - 0.05435(300) + 1.07822(2) = 19.30$ mpg.

12-7.

```
Predictor         Coef     SE Coef          T        P
Constant        383.80       36.22      10.60    0.002
X1             -3.6381      0.5665      -6.42    0.008
X2            -0.11168     0.04338      -2.57    0.082
S = 12.35       R-Sq = 98.5%     R-Sq(adj) = 97.5%

Analysis of Variance
Source            DF          SS          MS         F        P
Regression         2       29787       14894     97.59    0.002
Residual Error     3         458         153
Total              5       30245
```

a) $\hat{y} = 383.80 - 3.6381x_1 - 0.1119x_2$

b) $\hat{\sigma}^2 = 153.0$, $se(\hat{\beta}_0) = 36.22$, $se(\hat{\beta}_1) = 0.5665$, and $se(\hat{\beta}_2) = .04338$

c) $\hat{y} = 383.80 - 3.6381(25) - 0.1119(1000) = 180.95$

d)

```
Predictor         Coef     SE Coef          T        P
Constant         484.0       101.3       4.78    0.041
X1              -7.656       3.846      -1.99    0.185
X2             -0.2221      0.1129      -1.97    0.188
X1*X2         0.004087    0.003871       1.06    0.402
S = 12.12       R-Sq = 99.0%     R-Sq(adj) = 97.6%
Analysis of Variance
Source            DF          SS          MS         F        P
Regression         3     29951.4      9983.8     67.92    0.015
Residual Error     2       294.0       147.0
Total              5     30245.3
```

$$\hat{y} = 484.0 - 7.656x_1 - 0.222x_2 - 0.0041x_{12}$$

e) $\hat{\sigma}^2 = 147.0$, $se(\hat{\beta}_0) = 101.3$, $se(\hat{\beta}_1) = 3.846$, $se(\hat{\beta}_2) = 0.113$ and $se(\hat{\beta}_{12}) = 0.0039$

f) $\hat{y} = 484.0 - 7.656(25) - 0.222(1000) - 0.0041(25)(1000) = -31.3$

The predicted value is smaller

12-9.

```
Predictor        Coef     SE Coef        T        P
Constant        47.17       49.58     0.95    0.356
x1             -9.735       3.692    -2.64    0.018
x2             0.4283      0.2239     1.91    0.074
x3             18.237       1.312    13.90    0.000

S = 3.480       R-Sq = 99.4%     R-Sq(adj) = 99.3%

Analysis of Variance
Source          DF          SS          MS         F        P
Regression       3       30532       10177    840.55    0.000
Residual Error  16         194          12
Total           19       30725
```

a) $y = 47.174 - 9.7352x_1 + 0.4283x_2 + 18.2375x_3$

b) $\hat{\sigma}^2 = 12$

c) $se(\hat{\beta}_0) = 49.5815$, $se(\hat{\beta}_1) = 3.6916$, $se(\hat{\beta}_2) = 0.2239$, and $se(\hat{\beta}_3) = 1.312$

d) $y = 47.174 - 9.7352(14.5) + 0.4283(220) + 18.2375(5) = 91.43$

Section 12-2

12-13. n = 10, k = 2, p = 3, $\alpha = 0.05$

$H_0 : \beta_1 = \beta_2 = ... = \beta_k = 0$

$H_1 : \beta_j \neq 0$ for at least one j

$$SS_T = 371595.6 - \frac{(1916)^2}{10} = 4490$$

$$\mathbf{X'y} = \begin{bmatrix} \sum y_i \\ \sum x_{i1}y_i \\ \sum x_{i2}y_i \end{bmatrix} = \begin{bmatrix} 1030 \\ 21310 \\ 44174 \end{bmatrix}$$

$$\hat{\mathbf{a}}'\mathbf{X'y} = [171.054 \quad 3.713 \quad -1.126] \begin{bmatrix} 1916 \\ 43550.8 \\ 104736.8 \end{bmatrix} = 371535.9$$

$$SS_R = 371535.9 - \frac{1916^2}{10} = 4430.38$$

$$SS_E = SS_T - SS_R = 4490 - 4430.38 = 59.62$$

$$f_0 = \frac{\frac{SS_R}{k}}{\frac{SS_E}{n-p}} = \frac{4430.38/2}{59.62/7} = 260.09$$

$$f_{0.05,2,7} = 4.74$$

$$f_0 > f_{0.05,2,7}$$

Reject H_0 and conclude that the regression model is significant at $\alpha = 0.05$.

b) $H_0 : \beta_1 = 0$ $\qquad \beta_2 = 0$

$H_1 : \beta_1 \neq 0$ $\beta_2 \neq 0$

$t_0 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)}$ $t_0 = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)}$

$= \frac{3.713}{0.1934} = 19.20$ $= \frac{-1.126}{0.0861} = -13.08$

$t_{\alpha/2,7} = t_{.025,7} = 2.365$

Reject H_0 Reject H_0

Both regression coefficients are significant

12-17. a) $H_0 : \beta_1 = \beta_6 = 0$

$H_1 :$ at least one $\beta \neq 0$

$f_0 = 53.3162$

$f_{\alpha,2,22} = f_{.05,2,22} = 3.44$

$f_0 > f_{\alpha,2,22}$

Reject H_0 and conclude regression model is significant at $\alpha = 0.05$

b) $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$t_0 = -8.59$

$t_{.025,25-3} = t_{.025,22} = 2.074$

$|t_0| > t_{\alpha/2,22}$, Reject H_0 and conclude β_1 is significant at $\alpha = 0.05$

$H_0 : \beta_6 = 0$

$H_1 : \beta_6 \neq 0$

$\alpha = 0.05$

$t_0 = 1.5411$

$|t_0| \ngtr t_{\alpha/2,22}$, Do not reject H_0, conclude that evidence is not significant to state β_6 is significant at $\alpha = 0.05$.

No, only x_1 contributes significantly to the regression.

12-21. a) $H_0 : \beta_1 = \beta_2 = \beta_{12} = 0$

H_1 at least one $\beta_j \neq 0$

$\alpha = 0.05$

$f_0 = 67.92$

$f_{\alpha,3,2} = f_{.05,3,2} = 19.16$

$f_0 \ngtr f_{\alpha,3,2}$

Reject H_0

b) $H_0 : \beta_{12} = 0$

$H_1 : \beta_{12} \neq 0$

$\alpha = 0.05$

$SSR(\beta_{12} \mid \beta_1, \beta_2) = 29951.4 - 29787 = 164.4$

$f_0 = \frac{SSR}{MS_E} = \frac{164.4}{153} = 1.07$

$f_{.05,1,2} = 18.51$

$f_0 \not> f_{\alpha,1,2}$

Do not reject H_0

c) $\hat{\sigma}^2 = 147.0$

$\hat{\sigma}^2$ (no interaction term) = 153.0

$MS_E(\hat{\sigma}^2)$ was reduced in the interaction term model due to the addition of this term.

12-23. a) $H_0: \beta_1 = \beta_2 = \beta_3 = 0$ for all j

$H_1: \beta_j \neq 0$ for at least one j

$f_0 = 840.55$

$f_{.05,3,16} = 3.24$

$f_0 > f_{\alpha,3,16}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$

b) $\alpha = 0.05$ $t_{\alpha/2,n-p} = t_{.025,16} = 2.12$

$H_0: \beta_1 = 0$	$\beta_2 = 0$	$\beta_3 = 0$
$H_1: \beta_1 \neq 0$	$\beta_2 \neq 0$	$\beta_3 \neq 0$
$t_0 = -2.637$	$t_0 = 1.91$	$t_0 = 13.9$
$\lvert t_0 \rvert > t_{\alpha/2,16}$	$\lvert t_0 \rvert \not> t_{\alpha/2,16}$	$\lvert t_0 \rvert > t_{\alpha/2,16}$
Reject H_0	Do not reject H_0	Reject H_0

Sections 12-3 and 12-4

12-27. a) $-0.00657 \le \beta_8 \le -0.00122$

b) $\sqrt{\hat{\sigma}^2 \mathbf{x}_0' (\mathbf{X'X})^{-1} \mathbf{x}_0} = 0.497648 = se(\hat{\mu}_{Y|x_0})$

c) $\hat{\mu}_{Y|x_0} = -7.63449 + 0.00398(2000) + 0.24777(60) - 0.00389(1800) = 8.19$

$\hat{\mu}_{Y|x_0} \pm t_{.025,24}\ se(\hat{\mu}_{Y|x_0})$

$8.19 \pm (2.064)(0.497648)$

8.19 ± 1.03

$7.16 \le \mu_{Y|x_0} \le 9.22$

12-29. a) 95 % CI on coeficients

$\beta_1 \pm t_{a/2,n-p}(\hat{\beta}_1)$

$0.0972 \le \beta_1 \le 1.4174$

$-1.9646 \le \beta_2 \le 17.0026$

$-1.7953 \le \beta_3 \le 6.7613$

$-1.7941 \le \beta_4 \le 0.8319$

b) $\hat{\mu}_{Y|x_0} = 290.44$ $se(\hat{\mu}_{Y|x_0}) = 7.61$ $t_{.025,7} = 2.365$

$$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,n-p} \; se(\hat{\mu}_{Y|x_0})$$
$$290.44 \pm (2.365)(7.61)$$
$$272.44 \le \mu_{Y|x_0} \le 308.44$$

c) $\hat{y}_0 \pm t_{\alpha/2,n-p} \sqrt{\hat{\sigma}^2 (1 + \mathbf{x}_0' (\mathbf{X}'\mathbf{X})^{-1} \mathbf{x}_0)}$

$$290.44 \pm 2.365(14.038)$$
$$257.25 \le y_0 \le 323.64$$

12-31 a)95% Confidence Interval on coefficients

$$-0.595 \le \beta_2 \le 0.535$$
$$0.229 \le \beta_3 \le 0.812$$
$$-0.216 \le \beta_4 \le 0.013$$
$$-7.982 \le \beta_5 \le 2.977$$

b) $\hat{\mu}_{Y|x_0} = 8.99568$ $se(\hat{\mu}_{Y|x_0}) = 0.472445$ $t_{.025,14} = 2.145$

$$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,n-p} \; se(\hat{\mu}_{Y|x_0})$$
$$8.99568 \pm (2.145)(0.472445)$$
$$7.982 \le \mu_{Y|x_0} \le 10.009$$

c) $y_0 = 8\ 99568$ $se(\hat{y}_0) = 1.00121$

$$8.99568 \pm 2.145(1.00121)$$
$$6.8481 \le y_0 \le 11.143$$

12-35. a) $0.3882 \le \beta_{Pts} \le 0.5998$

b) $\hat{y} = -5.767703 + 0.496501 x_{Pts}$

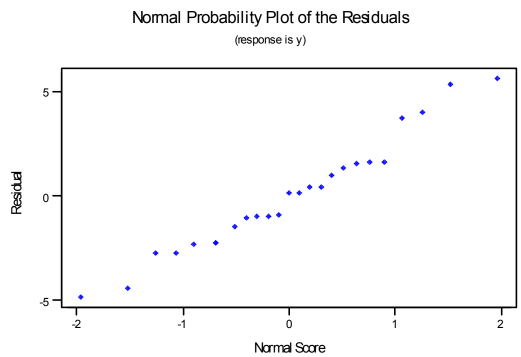
c) $0.4648 \le \beta_{Pts} \le 0.5282$

d) The simple linear regression model has the shorter interval. Yes, the simple linear regression model in this case is preferable.

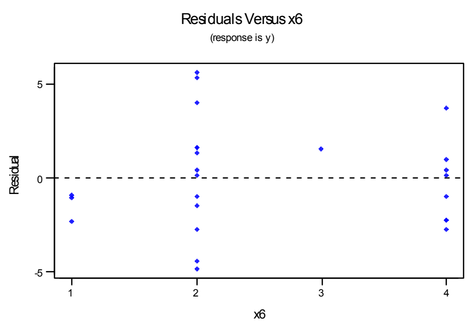
Section 12-5

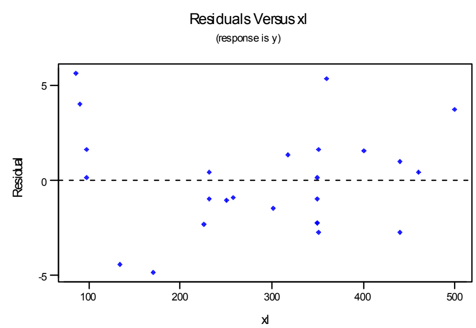
12-37. a) $r^2 = 0.82897$

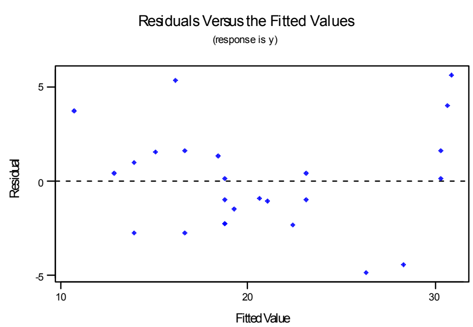
b) Normality assumption appears valid.



c) Assumption of constant variance appears reasonable.







d) Yes, observations 7, 10, and 18

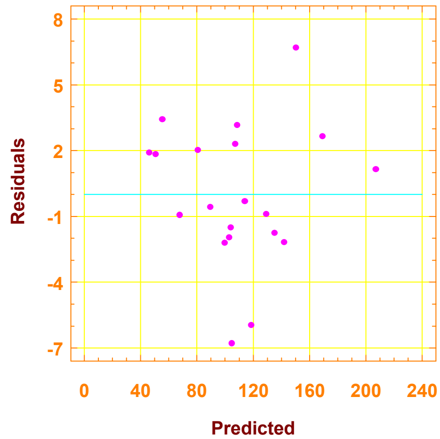
12-39. a) $r^2 = 0.985$

b) $r^2 = 0.990$

r^2 increases with addition of interaction term. No, adding additional regressor will always increase r^2

12-41 a) There is some indication of nonconstant variance since the residuals appear to "fan out" with increasing values of y.

Residual Plot for y



b)

Source	Sum of Squares	DF	Mean Square	F-Ratio	P-value
Model	30531.5	3	10177.2	840.546	.0000
Error	193.725	16	12.1078		
Total (Corr.)	30725.2	19			

R-squared = 0.993695 Stnd. error of est. = 3.47963
R-squared (Adj. for d.f.) = 0.992513 Durbin-Watson statistic = 1.77758

$R^2 = 0.9937$ or 99.37 %;

$R^2_{Adj} = 0.9925$ or 99.25%;

c)

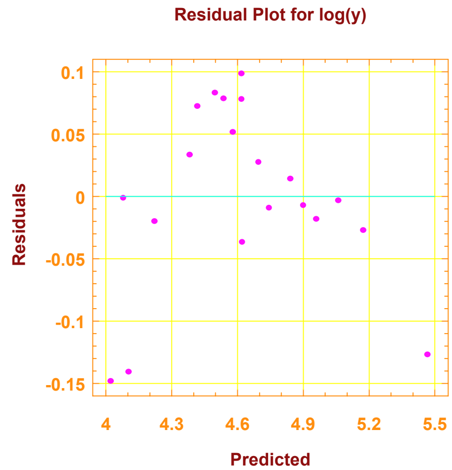
Model fitting results for: log(y)

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	6.22489	1.124522	5.5356	0.0000
x1	-0.16647	0.083727	-1.9882	0.0642
x2	-0.000228	0.005079	-0.0448	0.9648
x3	0.157312	0.029752	5.2875	0.0001

R-SQ. (ADJ.) = 0.9574 SE= 0.078919 MAE= 0.053775 DurbWat= 2.031
Previously: 0.0000 0.000000 0.000000 0.000
20 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

$$\hat{y}^* = 6.22489 - 0.16647x_1 - 0.000228x_2 + 0.157312x_3$$

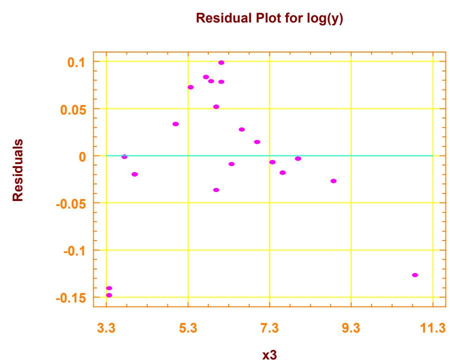
d)



Plot exhibits curvature

There is curvature in the plot. The plot does not give much more information as to which model is preferable.

e)



Plot exhibits curvature

Variance does not appear constant. Curvature is evident.

f)

Model fitting results for: log(y)

Independent variable	coefficient	std. error	t-value	sig.level
CONSTANT	6.222045	0.547157	11.3716	0.0000
x1	-0.198597	0.034022	-5.8374	0.0000
x2	0.009724	0.001864	5.2180	0.0001
1/x3	-4.436229	0.351293	-12.6283	0.0000

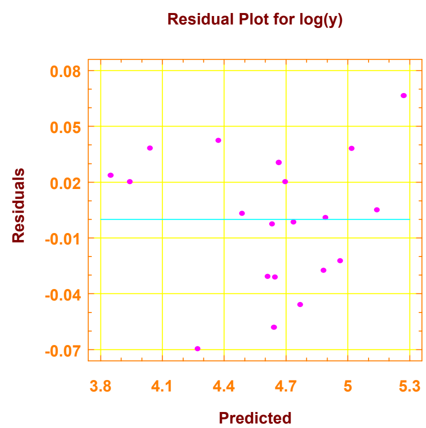
R-SQ. (ADJ.) = 0.9893	SE=	0.039499	MAE= 0.028896	DurbWat=	1.869
Previously: 0.9574		0.078919	0.053775		2.031

20 observations fitted, forecast(s) computed for 0 missing val. of dep. var.

Analysis of Variance for the Full Regression

Source	Sum of Squares	DF	Mean Square	F-Ratio	P-value
Model	2.75054	3	0.916847	587.649	.0000
Error	0.0249631	16	0.00156020		
Total (Corr.)	2.77550	19			

R-squared = 0.991006 Stnd. error of est. = 0.0394993
R-squared (Adj. for d.f.) = 0.98932 Durbin-Watson statistic = 1.86891



Using 1/x3

The residual plot indicates better conformance to assumptions.

Curvature is removed when using $1/x_3$ as the regressor instead of x_3 and the log of the response data.

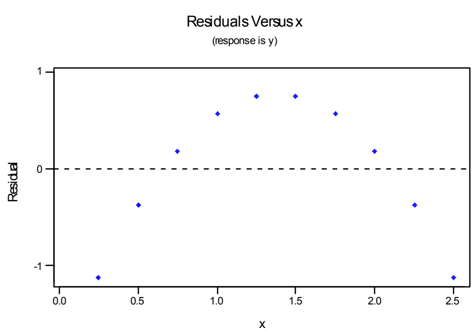
Section 12-6

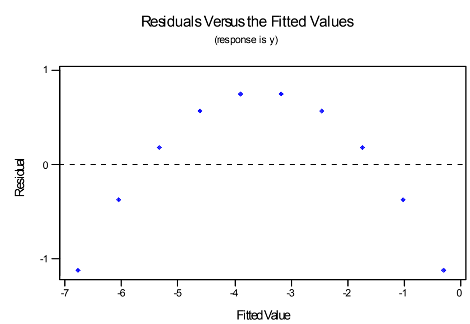
12-47. a) $\hat{y} = -1.633 + 1.232x - 1.495x^2$

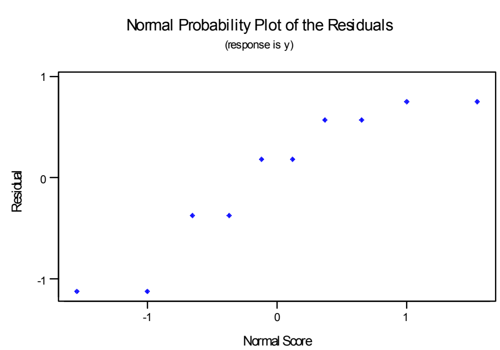
b) $f_0 = 1858613$, reject H_0

c) $t_0 = -601.64$, reject H_0

d) Model is acceptable, observation number 10 has large leverage.







12-49. $\hat{y} = 759.395 - 90.783x' - 47.166(x')^2$, where $x' = \dfrac{x - \bar{x}}{S_x}$

a) At $x = 285$ $x' = \dfrac{285 - 297.125}{11.9336} = -1.016$

$$\hat{y} = 759.395 - 90.783(-1.106) - 47.166(-1.106)^2 = 802.943 \text{ psi}$$

b) $\hat{y} = 759.395 - 90.783\left(\frac{x-297.125}{11.9336}\right) - 47.166\left(\frac{x-297.125}{11.9336}\right)^2$

$$\hat{y} = 759.395 - 7.607(x - 297.125) - 0.331(x - 297.125)^2$$

$$\hat{y} = -26204.14 + 189.09x - 0.331x^2$$

c) They are the same.

d) $\hat{y}' = 0.385 - 0.847x' - 0.440(x')^2$

where $y' = \dfrac{y - \bar{y}}{S_y}$ and $x' = \dfrac{x - \bar{x}}{S_x}$

The "proportion" of total variability explained is the same for both standardized and un-standardized models. Therefore, R^2 is the <u>same</u> for both models.

$y' = \beta_0^* + \beta_1^* x' + \beta_{11}^*(x')^2$ where $y' = \dfrac{y - \bar{y}}{S_y}$ and $x' = \dfrac{x - \bar{x}}{S_x}$ $y' = \beta_0^* + \beta_1^* x' + \beta_{11}^*(x')^2$

12-51 a)

Predictor	Coef	SE Coef	T	P
Constant	-1.769	1.287	-1.37	0.188
x1	0.4208	0.2942	1.43	0.172
x2	0.2225	0.1307	1.70	0.108
x3	-0.12800	0.07025	-1.82	0.087
x1x2	-0.01988	0.01204	-1.65	0.118
x1x3	0.009151	0.007621	1.20	0.247
x2x3	0.002576	0.007039	0.37	0.719
x1^2	-0.01932	0.01680	-1.15	0.267
x2^2	-0.00745	0.01205	-0.62	0.545
x3^3	0.000824	0.001441	0.57	0.575

S = 0.06092 R-Sq = 91.7% R-Sq(adj) = 87.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	9	0.655671	0.072852	19.63	0.000
Residual Error	16	0.059386	0.003712		
Total	25	0.715057			

$$\hat{y} = -1.769 + 0.421x_1 + 0.222x_2 - 0.128x_3 - 0.02x_1x_2 + 0.009x_1x_3 + 0.003x_2x_3 - 0.019x_1^2 - 0.007x_2^2 + 0.001x_3^2$$

b) H_0 : all $\beta_1 = \beta_2 = \beta_3 = \text{K} = \beta_{33} = 0$

H_1: at least 1 $\beta_j \neq 0$

$f_0 = 19.628$

$f_{.05,9,16} = 2.54$

$f_0 > f_{\alpha,9,16}$

Reject H_0 and conclude that the model is significant at $\alpha = 0.05$

c) Model is acceptable.

d) $H_0 : \beta_{11} = \beta_{22} = \beta_{33} = \beta_{12} = \beta_{13} = \beta_{23} = 0$

H_1: at least one $\beta_{ij} \neq 0$

$$f_0 = \frac{SS_R(\beta_{11}, \beta_{22}, \beta_{33}, \beta_{12}, \beta_{13}, \beta_{23} \mid \beta_1, \beta_2, \beta_3, \beta_0)/r}{MS_E} = \frac{\frac{0.0359}{6}}{0.003712} = 1.612$$

$$f_{.05,6,16} = 2.74$$

$$f_0 \not> f_{.05,6,16}$$

Do not reject H_0

$$SS_R(\beta_{12}, \beta_{13}, \beta_{23}, \beta_{11}, \beta_{22}, \beta_{33} \mid \beta_1, \beta_2, \beta_3, \beta_0) = SS_R(\beta_1, \beta_2, \beta_3, \beta_{12}, \beta_{13}, \beta_{23}, \beta_{11}, \beta_{22}, \beta_{33} \mid \beta_0) - SS_R(\beta_1\beta_2\beta_3 \mid \beta_0)$$

$$= 0.65567068 - 0.619763$$

$$= 0.0359$$

Reduced Model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

12-55. a) The min. MS_E equation is $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$

$MS_E = 6.58$ $\quad c_p = 5.88$

The min. C_p x_5, x_8, x_{10}

$C_p = 5.02$ $\quad MS_E = 7.97$

b) $\hat{y} = 34.434 - 0.048x_1$

$MS_E = 8.81$ $\quad C_p = 5.55$

c) Same as part b.

d) $\hat{y} = 0.341 + 2.862x_5 + 0.246x_8 - 0.010x_{10}$

$MS_E = 7.97$ $\quad C_p = 5.02$

e) Minimum C_p and backward elimination result in the same model. Stepwise and forward selection result in the same model. Because it is much smaller, the minimum C_p model seems preferable.

12-61. a) Min. C_p

$\hat{y} = -3.517 + 0.486x_1 - 0.156x_9$

$C_p = -1.67$

b) Min MS_E model is x_1, x_7, x_9, $MS_E = 1.67$, $C_p = -0.77$

$y = -5\ 964 + 0\ 495x_1 + 0\ 025x_7 - 0\ 163x_9$

c) Max. adjusted R^2 model is x_1, x_7, x_9, Adj. $R^2 = 0.98448$ Yes, same as Min. MS_E model.

Supplemental Exercises

12-65. a) $H_0: \beta_3^* = \beta_4 = \beta_5 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.01$

$f_0 = 1323.62$

$f_{.01,3,36} = 4.38$

$f_0 >> f_{\alpha,3,36}$

Reject H_0 and conclude regression is significant.

P-value < 0.00001

b) $\alpha = 0.01$ $t_{.005,36} = 2.72$

$H_0: \beta_3^* = 0$	$H_0: \beta_4 = 0$	$H_0: \beta_5 = 0$
$H_1: \beta_3^* \neq 0$	$H_1: \beta_4 \neq 0$	$H_1: \beta_5 \neq 0$
$t_0 = -1.32$	$t_0 = 19.97$	$t_0 = 2.48$
$\lvert t_0 \rvert \not> t_{\alpha/2,36}$	$\lvert t_0 \rvert > t_{\alpha/2,36}$	$\lvert t_0 \rvert \not> t_{\alpha/2,36}$
Do not reject H_0	Reject H_0	Do not reject H_0

Only regressor x_4 is significant

c) Curvature is evident in the residuals vs. regressor plots from this model.

12-67. a) $\hat{y} = -0.908 + 5.482x_1^* + 1.126x_2^* - 3.920x_3^* - 1.143x_4^*$

b) $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

$H_1: \beta_j \neq 0$ for at least one j

$\alpha = 0.05$

$f_0 = 109.02$

$f_{.05,4,19} = 2.90$

$f_0 >> f_{\alpha,4,19}$

Reject H_0 and conclude regression is significant at $\alpha = 0.05$.

$\alpha = 0.05$ $t_{.025,19} = 2.093$

$H_0: \beta_1 = 0$	$H_0: \beta_2 = 0$	$H_0: \beta_3 = 0$	$H_0: \beta_4 = 0$
$H_1: \beta_1 \neq 0$	$H_1: \beta_2 \neq 0$	$H_1: \beta_3 \neq 0$	$H_1: \beta_4 \neq 0$
$t_0 = 11.27$	$t_0 = 14.59$	$t_0 = -6.98$	$t_0 = -8.11$
$\lvert t_0 \rvert > t_{\alpha/2,19}$	$\lvert t_0 \rvert > t_{\alpha/2,19}$	$\lvert t_0 \rvert > t_{\alpha/2,19}$	$\lvert t_0 \rvert > t_{\alpha/2,19}$
Reject H_0	Reject H_0	Reject H_0	Reject H_0

c) The residual plots are more pleasing than those in Exercise 12-66.

12-69. a) $\hat{y} = -3982.1 + 1.0964x_1 + 0.1843x_3 + 3.7456x_4 + 0.8343x_5 - 16.2781x_6$

$MS_E(p) = 694.93$ $C_p = 5.62$

b) $\hat{y} = -4280.2 + 1.442x_1 + 0.209x_3 + 0.6467x_5 - 17.5103x_6$

$MS_E(p) = 714.20$ $C_p = 5.57$

c) Same as model b.

d) Models from parts b. and c. are identical. Model in part a. is the same with x_4 added in.

MS_E model in part a. = 694.93 $C_p = 5.62$

MS_E model in parts b.&c. = 714.20 $C_p = 5.57$

12-71. a) $VIF(\hat{\beta}_3^*) = 51.86$

$VIF(\hat{\beta}_4) = 9.11$

$VIF(\hat{\beta}_5) = 28.99$

Yes, VIFs for x_3^* and x_5 exceed 10.

b) Model from Exercise 12-65: $\hat{y} = 19.69 - 1.27x_3^* + 0.005x_4 + 0.0004x_5$

12-73. a) $R^2 = \frac{SS_R}{SS_T}$

$$SS_R = R^2(SS_T) = 0.94(0.50) = 0.47$$
$$SS_E = SS_T - SS_R = 0.5 - 0.47 = 0.03$$
$$H_0: \beta_1 = \beta_2 = ... = \beta_6 = 0$$
$H_1: \beta_j \neq 0$ for at least one *j*.

$\alpha = 0.05$

$$f_0 = \frac{SS_R / k}{SS_E / n - p} = \frac{0.47/6}{0.03/7} = 18.28$$
$$f_{.05,6,7} = 3.87$$
$$f_0 > f_{\alpha,6,7}$$
Reject H_0.

b) $k = 5$ $n = 14$ $p = 6$ $R^2 = 0.92$

$$SS_R' = R^2(SS_T) = 0.92(0.50) = 0.46$$
$$SS_E' = SS_T - SS_R' = 0.5 - 0.46 = 0.04$$
$$SS_R(\beta_j, \beta_{i,i=1,2,\Lambda .6,i\neq j} \mid \beta_0) = SS_R(full) - SS_R(reduced) = 0.47 - 0.46 = 0.01$$
$$f_0 = \frac{SS_R(\beta_j \mid \beta_{i,i=1,2,\Lambda ,6,i\neq j} \mid \beta_0)/r}{SS_E'/(n-p)} = \frac{0.01/1}{0.04/8} = 2$$
$$f_{.05,1,8} = 5.32$$
$$f_0 \not> f_{\alpha,1,8}$$
Do not reject H_0 and conclude that the evidence is insufficient to claim that the removed variable is significant at $\alpha = 0.05$

c) $MS_E(reduced) = \frac{SS_E}{n-p} = \frac{0.04}{8} = 0.005$

$$MS_E(full) = \frac{0.03}{7} = 0.00429$$
No, the MS_E is larger for the reduced model, although not by much. Generally, if adding a variable to a model reduces the MS_E it is an indication that the variable may be useful in explaining the response variable. Here the decrease in MS_E is not very great because the added variable had no real explanatory power.