Chapter 14

14–1. (a) $\hat{y} = 10.5 - 0.00156$ yards

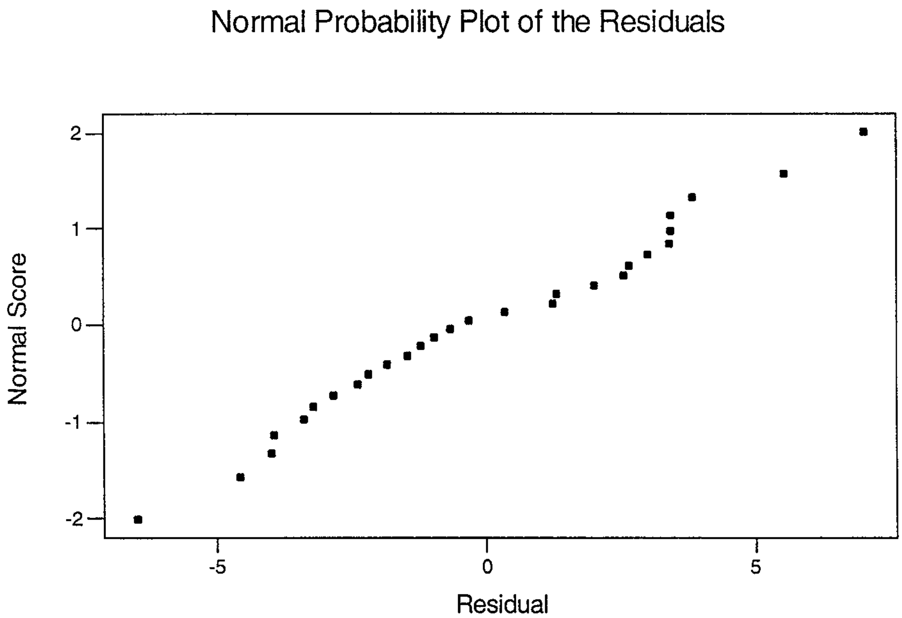
(b)

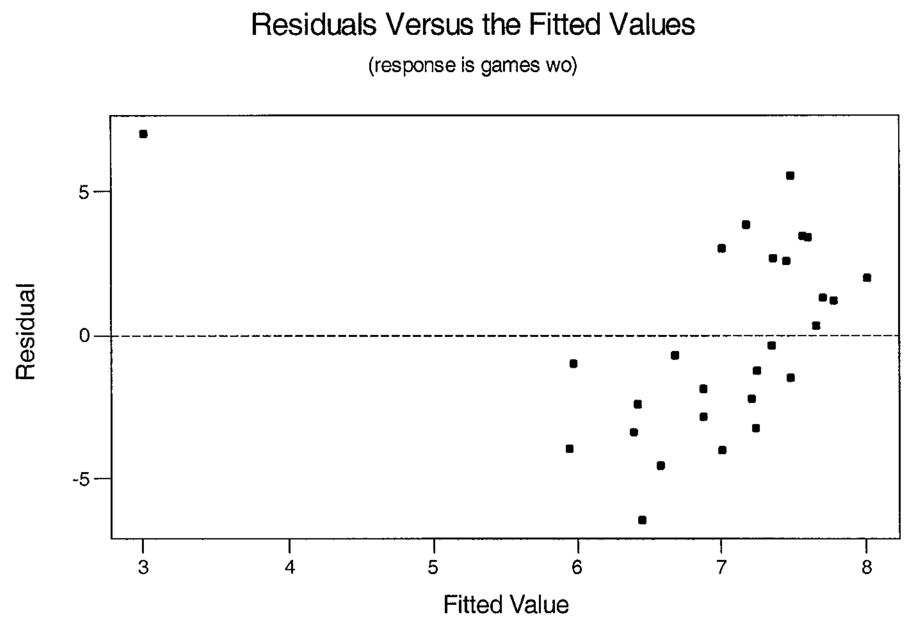
```
Predictor        Coef      SE Coef        T        P
Constant       10.451        2.514     4.16    0.000
yards       -0.001565     0.001090    -1.43    0.163

S = 3.414       R-Sq = 7.3%      R-Sq(adj) = 3.8%

Analysis of Variance
Source           DF           SS        MS          F        P
Regression        1        23.99     23.99       2.06    0.163
Residual Error   26       302.97     11.65
Total            27       326.96
```

(c) $-0.00380 \leq \beta_1 \leq 0.00068$

(d) $R^2 = 7.3\%$

(e)





Based on the residual plots there appears to be a severe outlier. This point should be investigated and if necessary, the point removed and the analysis rerun.

14–2. $\hat{y} = 10.5 - 0.00156(1800) = 7.63$ or approximately 8.
$6 \leq E(y|x_0 = 1800) \leq 9.27$

14–3. (a) $\hat{y} = 31.6 - 0.0410$ displacement

(b)

Predictor	Coef	SE Coef	T	P
Constant	31.648	1.812	17.46	0.000
displace	-0.040975	0.005406	-7.58	0.000

S = 1.976   R-Sq = 81.5%   R-Sq(adj) = 80.1%$

Analysis of Variance

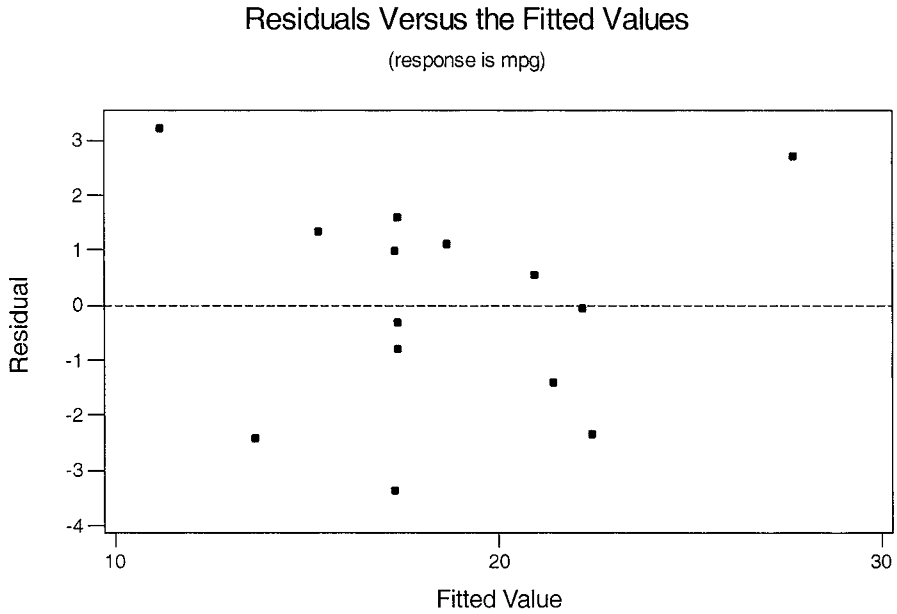
Source	DF	SS	MS	F	P
Regression	1	224.43	224.43	57.45	0.000
Residual Error	13	50.79	3.91		
Total	14	275.21			

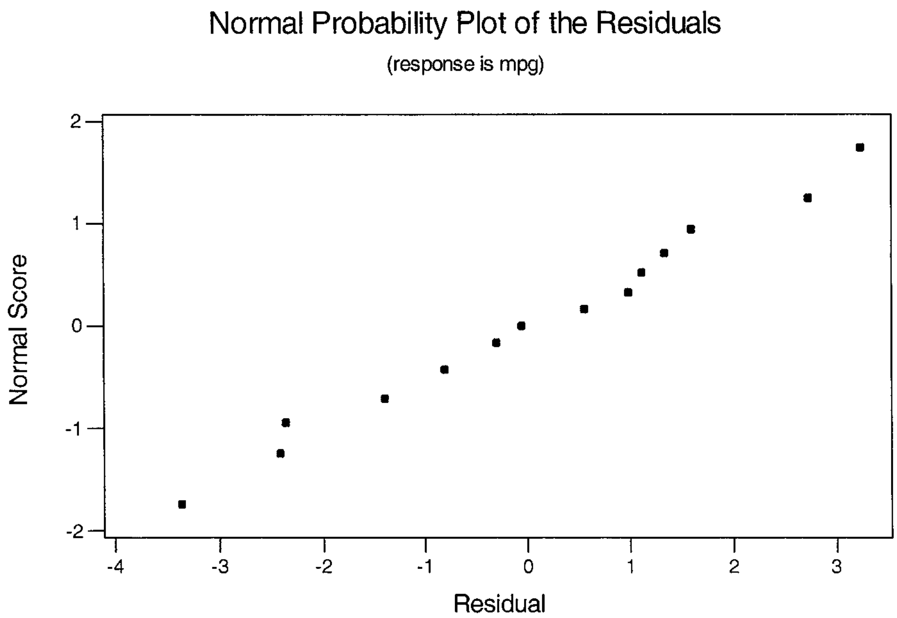
(c) $R^2 = 81.5\%$

(d) $\hat{y} = 20.381$, $19.374 \leq E(y|x_0 = 275) \leq 21.388$

14–4. $\hat{y} = 31.6 - 0.0410(275) = 20.381$, $10.37 \leq E(y|x_0 = 275) \leq 21.39$

14–5.





14–6. (a) $\hat{y} = 13.3 + 3.32$ taxes

(b)

Predictor	Coef	SE Coef	T	P
Constant	13.320	2.572	5.18	0.000
taxes	3.3244	0.3903	8.52	0.000

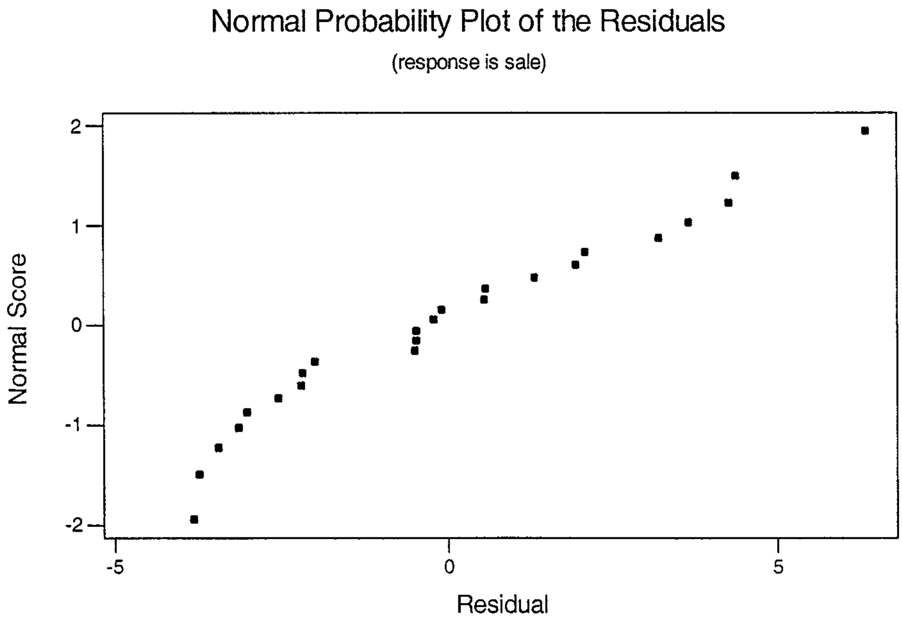
S = 2.961   R-Sq = 76.7%   R-Sq(adj) = 75.7%

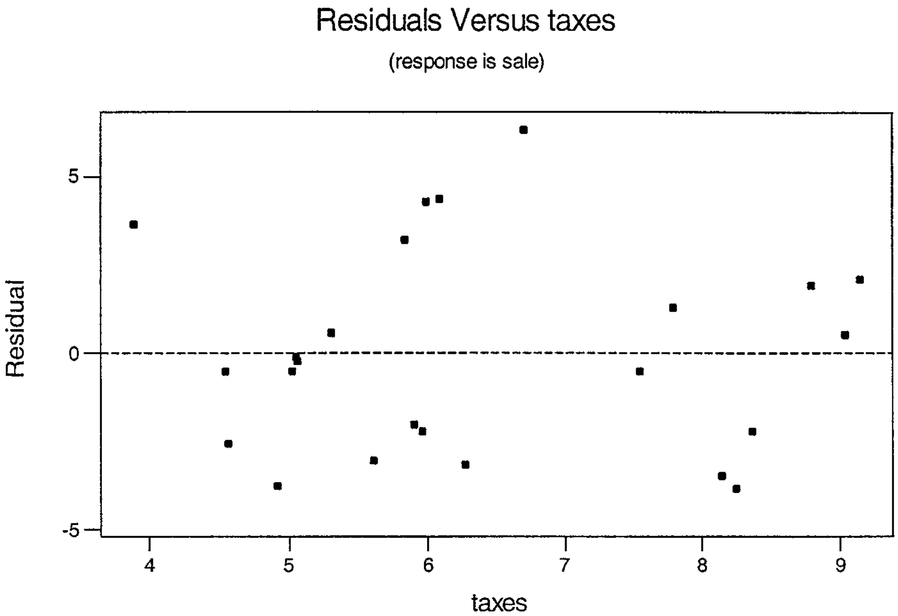
Analysis of Variance

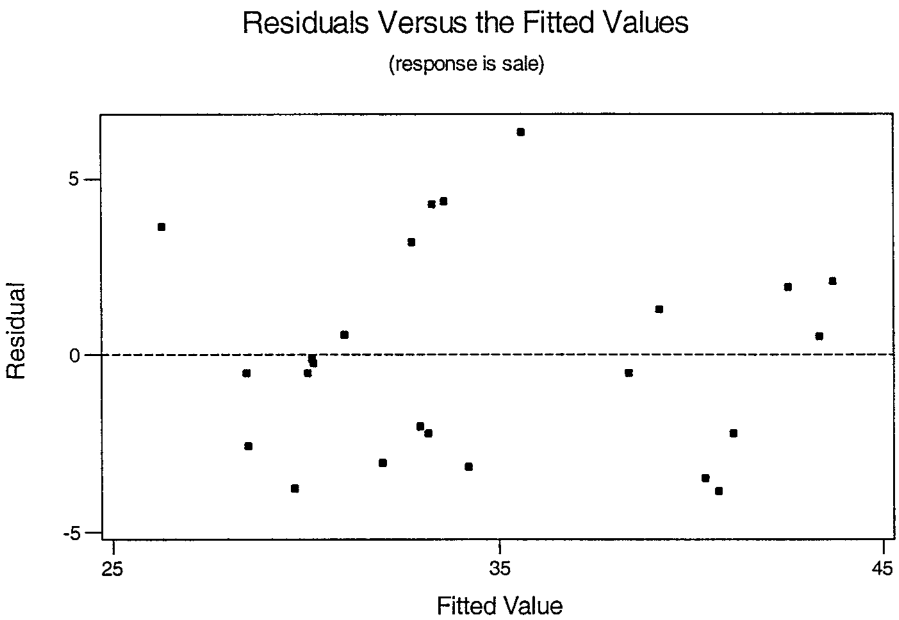
Source	DF	SS	MS	F	P
Regression	1	636.16	636.16	72.56	0.000
Residual Error	22	192.89	8.77		
Total	23	829.05			

(c) 76.7%

(d)




Residuals Versus taxes
(response is sale)
Residual
5
0
-5
4
5
6
7
8
9
taxes


Residuals Versus the Fitted Values
(response is sale)
Residual
5
0
-5
25
35
45
Fitted Value

14–7. (a) $\hat{y} = 93.3 + 15.6x$

(b)

Predictor	Coef	SE Coef	T	P
Constant	93.34	10.51	8.88	0.000
x	15.649	4.345	3.60	0.003

```
S = 11.63       R-Sq = 48.1%     R-Sq(adj) = 44.4%
```

Analysis of Variance

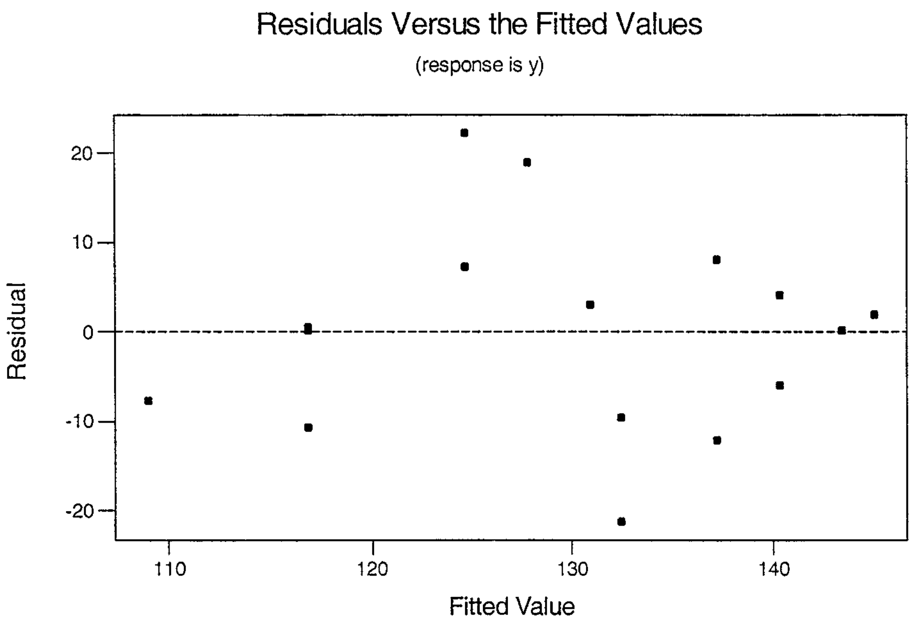
Source	DF	SS	MS	F	P
Regression	1	1755.8	1755.8	12.97	0.003
Residual Error	14	1895.0	135.4		
Lack of Fit	8	1378.6	172.3	2.00	0.207
Pure Error	6	516.4	86.1		
Total	15	3650.8			

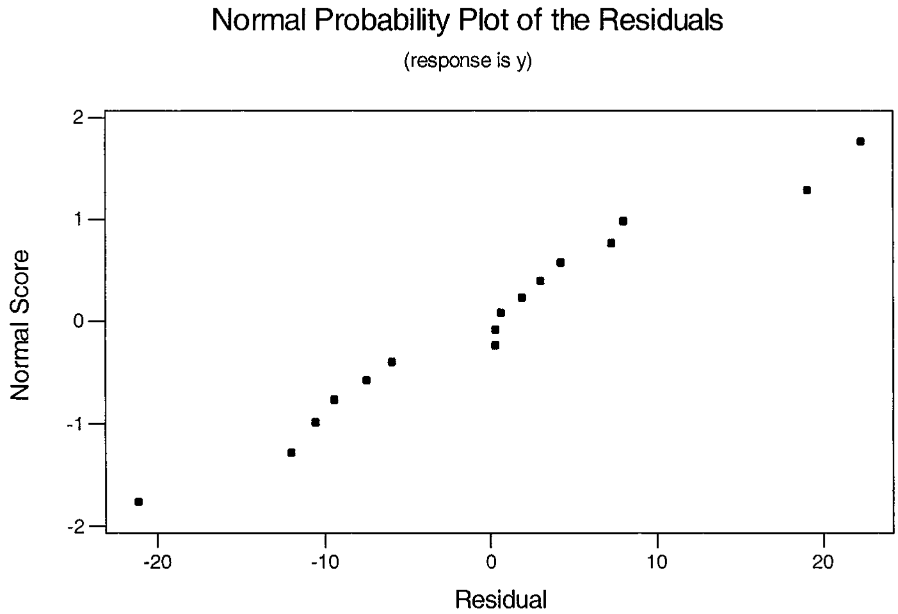
(c) $7.997 \leq \beta_1 \leq 23.299$

(d) $74.828 \leq \beta_0 \leq 111.852$

(e) $(126.012, 138.910)$

14–8.





14–9. (a) $\hat{y} = -6.34 + 9.21$ temp

(b)

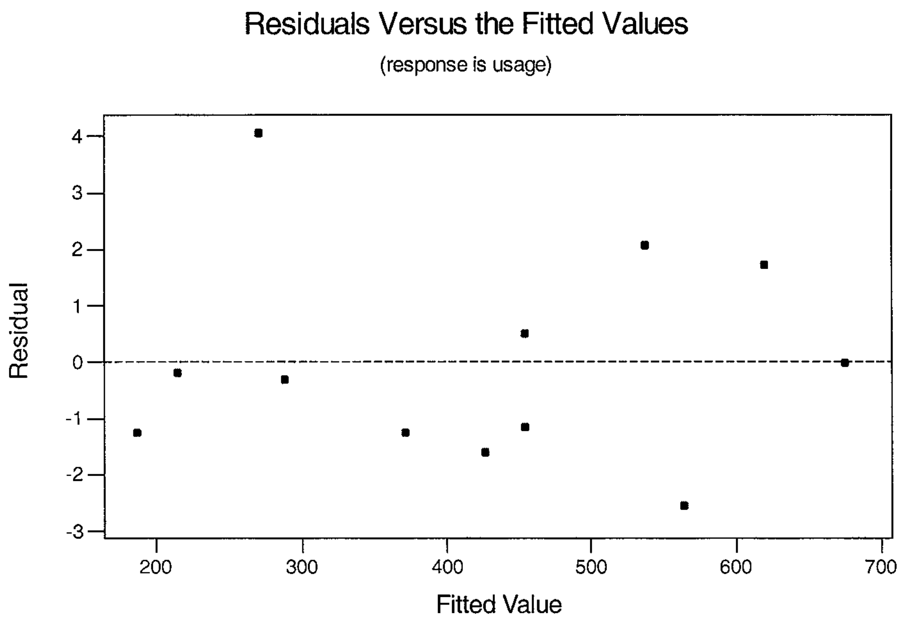
```
Predictor         Coef      SE Coef          T          P
Constant        -6.336        1.668      -3.80      0.003
temp           9.20836      0.03377     272.64      0.000

S = 1.943       R-Sq = 100.0%      R-Sq(adj) = 100.0%

Analysis of Variance
Source            DF           SS          MS          F         P
Regression         1       280583      280583   74334.36     0.000
Residual Error    10           38           4
Total             11       280621
```

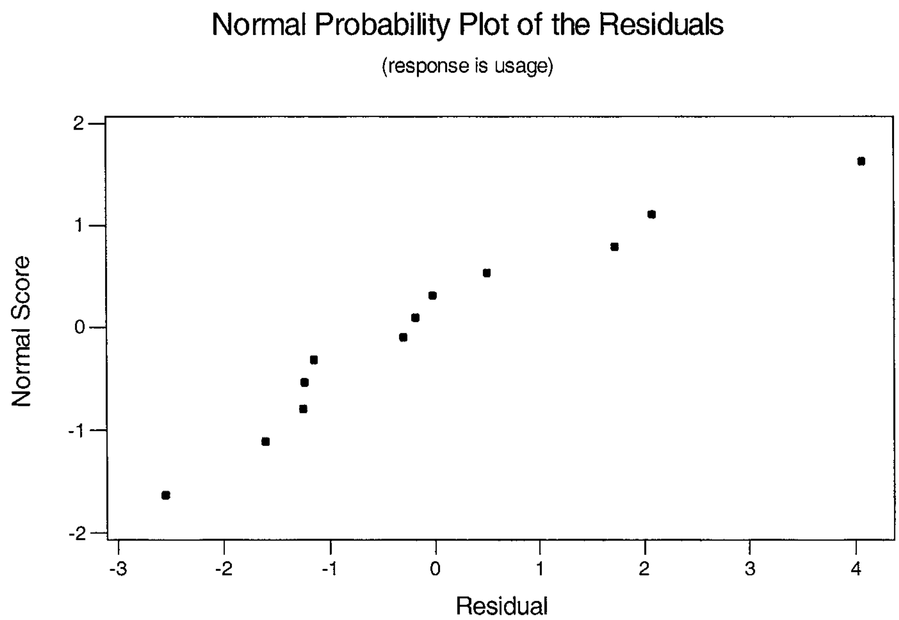
(c) $t_0 = \dfrac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\dfrac{MS_E}{S_{XX}}}} = \dfrac{9.20836 - 10}{\sqrt{\dfrac{4}{3309}}} = -23.41$; $t_{0.025,10} = 2.228$; Reject $\beta_1 = 0$

(d) $525.58 \leq E(y|X = 50) \leq 529.91$

(e) $521.22 \leq y_{x=58} \leq 534.28$

14–10.


Residuals Versus the Fitted Values
(response is usage)
Residual
4
3
2
1
0
-1
-2
-3
200
300
400
500
600
700
Fitted Value


Normal Probability Plot of the Residuals
(response is usage)
Normal Score
2
1
0
-1
-2
-3
-2
-1
0
1
2
3
4
Residual

14–11. (a) $\hat{y} = 77.7895 + 11.8634x$

(b)

Predictor	Coef	SE Coef	T	P
Constant	77.863	4.199	18.54	0.000
hydrocar	11.801	3.485	3.39	0.003

S = 3.597 R-Sq = 38.9% R-Sq(adj) = 35.5%

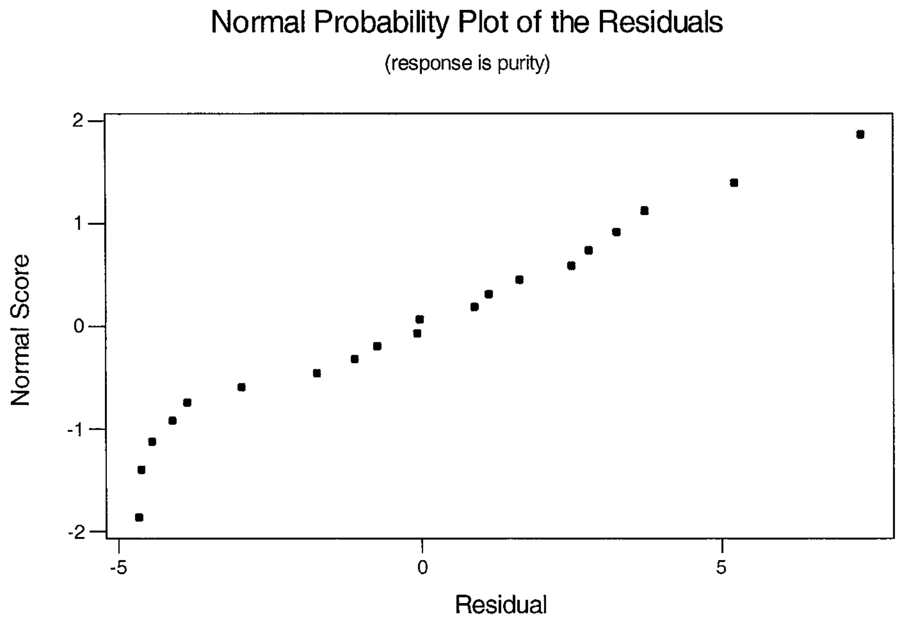
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	148.31	148.31	11.47	0.003
Residual Error	18	232.83	12.94		

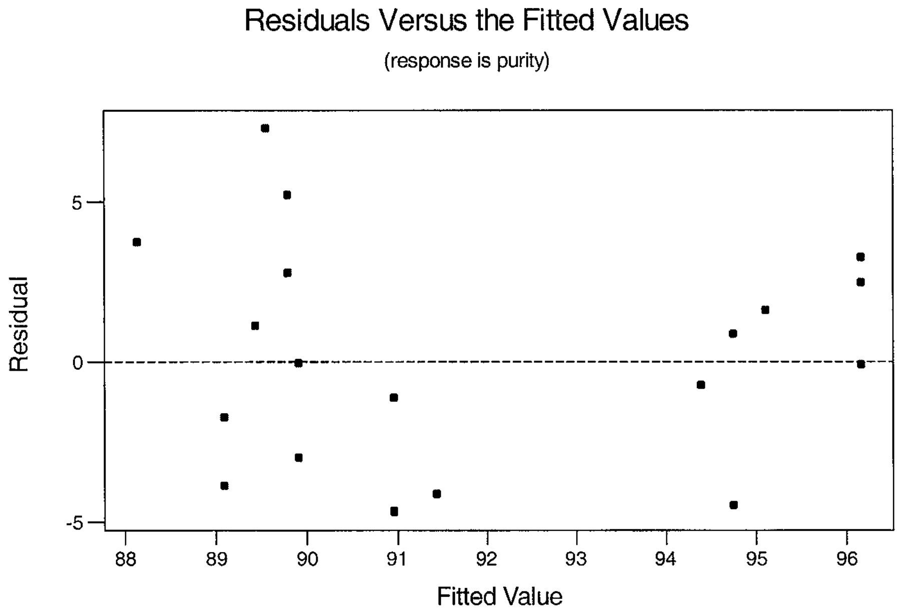
(c) 38.9%

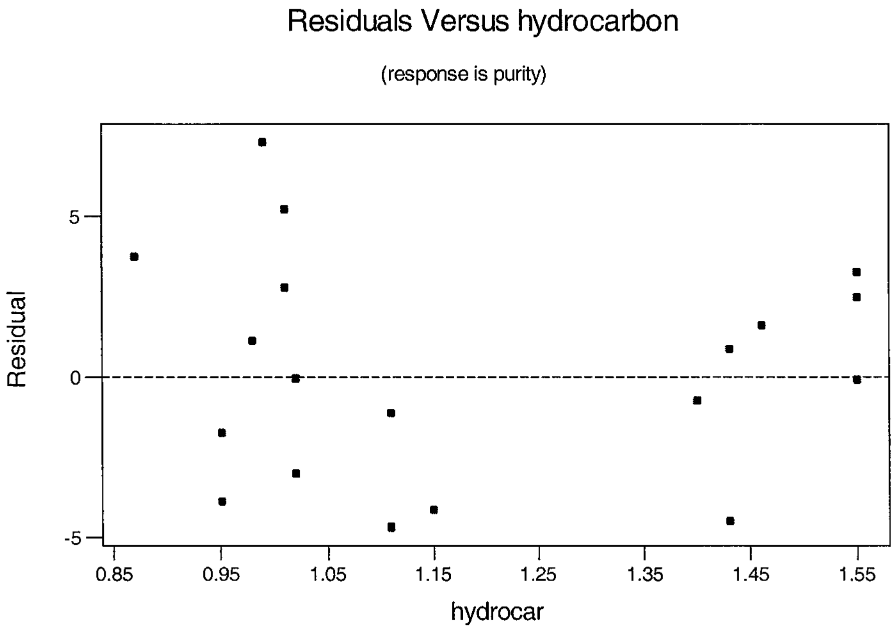
(d) $4.5661 \leq \beta \leq 19.1607$

14–12. (a)



(b)


Residuals Versus the Fitted Values
(response is purity)
Residual
5
0
-5
88
89
90
91
92
93
94
95
96
Fitted Value


Residuals Versus hydrocarbon
(response is purity)
Residual
5
0
-5
0.85
0.95
1.05
1.15
1.25
1.35
1.45
1.55
hydrocar

14–13. (a) $AvgSize = -1922.7 + 564.54$ Level

(b)

Predictor	Coef	SE Coef	T	P
Constant	-1922.7	530.9	-3.62	0.003
Level	564.54	32.74	17.24	0.000

S = 459.0 R-Sq = 95.5% R-Sq(adj) = 95.2%

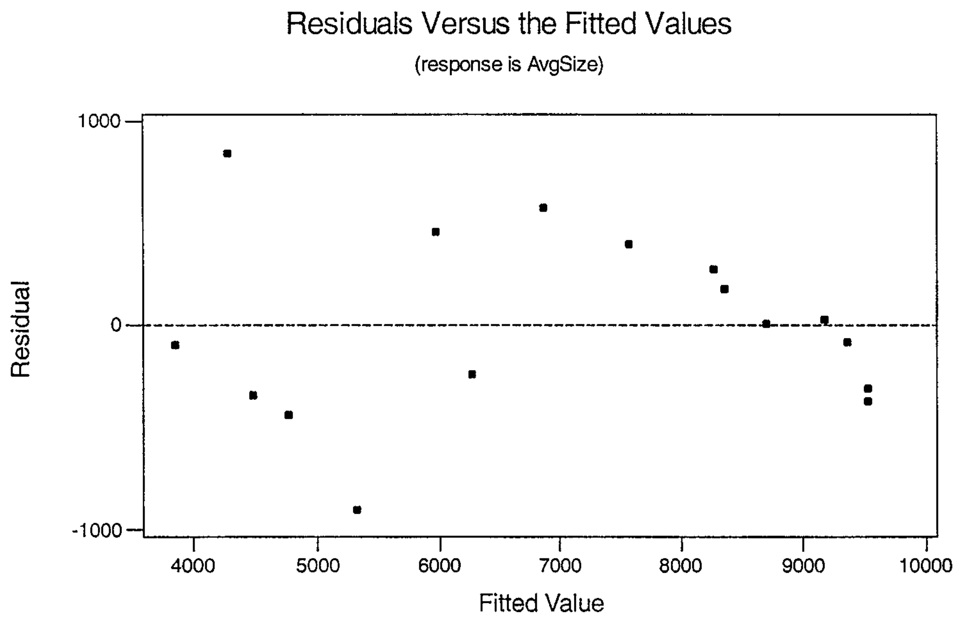
Analysis of Variance

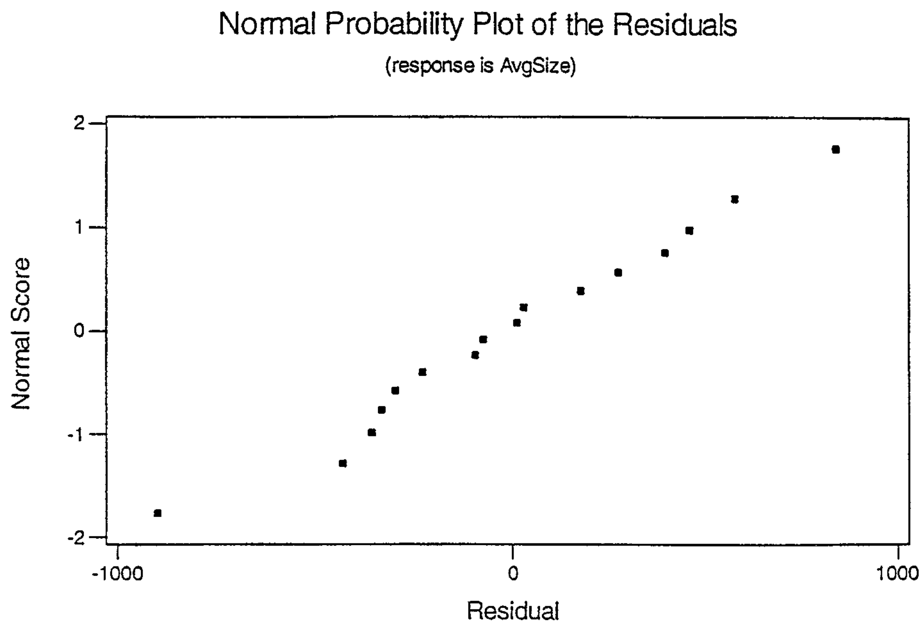
Source	DF	SS	MS	F	P
Regression	1	62660784	62660784	297.38	0.000
Residual Error	14	2949914	210708		
Total	15	65610698			

(c) $(0.0015, 0.0019)$

(d) 95.5%

(e)





14–14. x = OR Grade, y = Statistics Grade

(a) $\hat{y} = -0.0280 + 0.9910x$

(b) $r = 0.9033$

(c) $t_0 = \dfrac{r\sqrt{n-2}}{r\sqrt{1-r^2}} = \dfrac{0.9033\sqrt{18}}{\sqrt{1-0.8160}} = 8.93$, reject H_0.

(d) $Z_0 = (\text{arctanh}(0.9033) - \text{arctanh}(0.5))\sqrt{17} = 3.88$, reject H_0.

(e) $0.7676 \le \rho \le 0.9615$

14–15. x = weight, y = BP

(a) $\hat{y} = 69.1044 + 0.4194x$

(b) $r = 0.7735$

(c) $t_0 = \dfrac{r\sqrt{n-3}}{\sqrt{1-r^2}} = \dfrac{0.7735\sqrt{23}}{\sqrt{1-0.5983}} = 5.85$, reject H_0.

(d) $Z_0 = (\text{arctanh}(0.7735) - \text{arctanh}(0.6))\sqrt{23} = 1.61$, do not reject

(e) $0.5513 \le \rho \le 0.8932$

14–16. Note that $SS_R = \hat{\beta}_1 S_{xy} = \hat{\beta}_1^2 S_{xx}$, and

$$V(\hat{\beta}_1) = E(\hat{\beta}_1^2) - [E(\hat{\beta}_1)]^2 = \frac{\sigma^2}{S_{xx}}$$

$$E(\hat{\beta}_1^2) = \beta_1^2 + \frac{\sigma^2}{S_{xx}}$$

Therefore

$$E(SS_R) = E(\hat{\beta}_1^2)S_{xx} = \sigma^2 + \beta_1^2 S_{xx}$$

$$E(MS_R) = E\left(\frac{SS_R}{1}\right) = \sigma^2 + \beta_1^2 S_{xx}$$

14–17. $E(\hat{\beta}_1) = E\left(\frac{S_{xy}}{S_{xx}}\right) = \frac{1}{S_{xx}}E(S_{xy})$

$$= \frac{1}{S_{xx}} E \sum_{i=1}^{n} (x_{i1} - \overline{x}_1) y_i$$

$$= \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_{i1} - \overline{x}_1) E(y_i)$$

$$= \frac{1}{S_{xx}} \sum_{i=1}^{n} (x_{i1} - \overline{x}_1)(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})$$

$$= \beta_1 + \beta_2 \frac{\sum_{i=1}^{n} x_{i2}(x_{i1} - \overline{x}_1)}{S_{xx}}$$

In general, $\hat{\beta}_1$ is a biased estimator of β_1.

14–18. $V(\hat{\beta}_1) = \sigma^2/S_{xx}$, which is minimized if we can make S_{xx} is as large as possible. Since $\sum_{i=1}^{n}(x_i - \overline{x})^2 = S_{xx}$, place the x's as far from $\overline{x}$ as possible. If n is even, put $n/2$ trials at each end of the region of interest. If n is odd, put 1 trial at the center and $(n-1)/2$ trials at each end. These designs should be used only when you are positive that the relationship between y and x is linear.

14–19. $$L = \sum_{i=1}^{n} w_i(y_i - \beta_0 - \beta_1 x_i)^2$$

$$\frac{\partial L}{\partial \beta_0} = -2\sum_{i=1}^{n} w_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0.$$

$$\frac{\partial L}{\partial \beta_1} = -2\sum_{i=1}^{n} w_i x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0.$$

Simplification of these two equations gives the normal equations for weighted least squares.

14–20. If $y = (\beta_0 + \beta_1 x + \epsilon)^{-1}$, then $1/y = \beta_0 + \beta_1 x + \epsilon$ is a straight-line regression model. The scatter diagram of $y^* = 1/y$ versus x is linear.

x	10	15	18	12	9	8	11	6
y^*	5.88	7.69	11.11	6.67	5.00	4.76	5.56	4.17

14–21. The no-intercept model would be the form $y = \beta_1 x + \epsilon$. This model is likely not a good one for this problem, because there is no data near the point $x = 0$, $y = 0$, and it is probably unwise to extrapolate the linear relationship back to the origin. The intercept is often just a parameter that improves the fit of the model to the data in the region where the data were collected.

14–22. $b = 14$ $\quad \Sigma x_i = 65.262$ $\quad \Sigma x_i^2 = 385.194$ $\quad \overline{x} = 4.662$

$\Sigma y_i = 208$ $\quad \Sigma y_i^2 = 3490$ $\quad \overline{y} = 14.857$ $\quad \Sigma x_i y_i = 1148.08$

$S_{xx} = 80.989$ $\quad S_{xy} = 178.473$ $\quad S_{yy} = 599.714$

$$\hat{\beta}_1 = 2.204$$
$$\hat{\beta}_0 = 4.582$$

$\hat{y} = 4.582 + 2.204x$

$$r = \frac{S_{xy}}{(S_{xx}S_{yy})^{1/2}} = 0.9919, \quad R^2 = 0.9839$$

$H_0\colon\ \rho = 0$ $\qquad t_0 = 27.08 > t_{0.05,12} = 1.782.$ $\quad \therefore$ reject H_0

$H_1\colon\ \rho \neq 0$

A strong correlation does not imply a cause and effect relationship.