3.2 **Finding $\boldsymbol{\beta_0}$ and $\boldsymbol{\beta_1}$.** The equation for a straight line (deterministic) is $y = \beta_0 + \beta_1 x$. If the line passes through the point (0, 1), then $x = 0, y = 1$ must satisfy the equation. That is, $1 = \beta_0 + \beta_1(0)$. Similarly, if the line passes through the point (2, 3), then $x = 2, y = 3$ must satisfy the equation: $3 = \beta_0 + \beta_1(2)$. Use these two equations to solve for β_0 and β_1, and find the equation of the line that passes through the points (0, 1) and (2, 3).

$$\beta_0 = 1 - 0 = 1, \beta_1 = 3 - 1 = 2, \therefore y = 1 + 2x$$

Ans: $\boldsymbol{y = 1 + 2x}$

3.6 **Learning the mechanics.** Use the method of least squares to fit a straight line to these six data points:

x	1	2	3	4	5	6
y	1	2	2	3	5	5

(a) What is the least squares estimates of β_0 and β_1?

$$\sum x = 1 + 2 + 3 + 4 + 5 + 6 = 21, \sum y = 1 + 2 + 2 + 3 + 5 + 5 = 18$$

$$\sum x^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 = 1 + 4 + 9 + 16 + 25 + 36 = 91$$

$$\sum y^2 = 1^2 + 2^2 + 2^2 + 3^2 + 5^2 + 5^2 = 1 + 4 + 4 + 9 + 25 + 25 = 68$$

$$\bar{x} = 21 \div 6 = 3.5, \bar{y} = 18 \div 6 = 3$$

$$\sum xy = 1(1) + 2(2) + 3(2) + 4(3) + 5(5) + 6(5) = 1 + 4 + 6 + 12 + 25 + 30 = 78$$

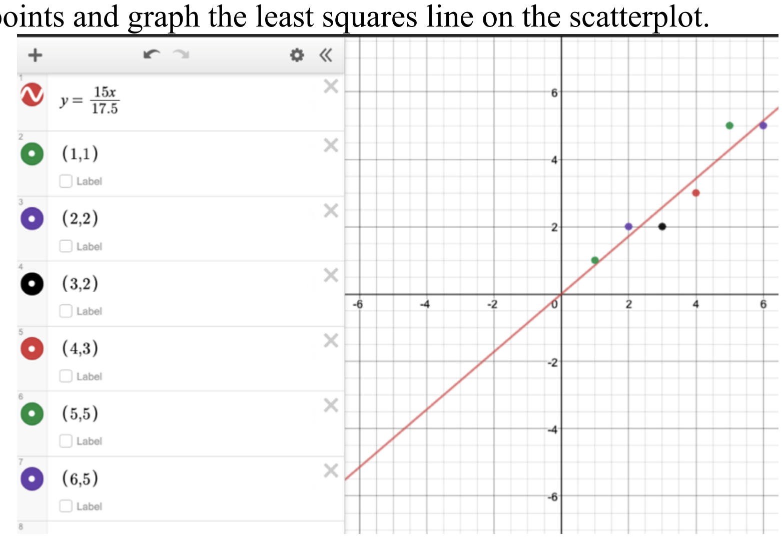
$$SS_{xx} = \sum x^2 - n(\bar{x})^2 = 91 - 6(3.5)^2 = 91 - 73.5 = 17.5$$

$$SS_{xy} = \sum xy - n\overline{xy} = 78 - 6(3.5)(3) = 78 - 63 = 15$$

$$\hat{\beta}_1 = \frac{SS_{xx}}{SS_{xy}} = \frac{15}{17.5} = .8571, \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 3 - (.8571)(3.5) = 0$$

Ans: $\boldsymbol{\hat{\beta}_0 = 0, \hat{\beta}_1 = .8571}$

(b) Plot the data points and graph the least squares line on the scatterplot.



3.8 **Predicting home sales price.** Real estate investors, homebuyers, and homeowners often use the appraised (or market) value of a property as a basis for predicting sales price. Data on sale prices and total appraised values of 76 residential properties sold in 2008 in an upscale Tampa, Florida, neighborhood named Tampa Palms are saved in the TAMPALMS file. The first five and last five observations of the data set are listed in the accompanying table.

TAMPALMS

PROPERTY	MARKET VALUE (THOUS.)	SALE PRICE (THOUS.)
1	\$184.44	\$382.0
2	191.00	230.0
3	159.83	220.0
4	189.22	277.0
5	151.61	205.0
⋮	⋮	⋮
72	263.40	325.0
73	194.58	252.0
74	219.15	270.0
75	322.67	305.0
76	325.96	450.0

Source: Hillsborough County (Florida) Property Appraiser's Office.

(a) Propose a straight-line model to relate the appraised property value x to the sale price y for residential properties in this neighborhood.
The straight-line model is $y = \beta_0 + \beta_1 x + c$

Ans: $\boldsymbol{y = \beta_0 + \beta_1 x + c}$

(b) A MINITAB scatterplot of the data is shown on the previous page. [*Note*: Both sale price and total market value are shown in thousands of dollars.] Does it appear that a straight-line model will be an appropriate fit to the data?

MINITAB Output for Exercise 3.8

Yes. The data form a rather straight line from the lower left of the plot to the upper right.

Ans: Yes

(c) A MINITAB simple linear regression printout is also shown (p. 100). Find the equation of the best-fitting line through the data on the printout.
The fitted model is $\hat{y} = 1.36 + 1.40827x$

Ans: $\boldsymbol{\hat{y} = 1.36 + 1.40827x}$

```
The regression equation is
Sale_Price = 1.4 + 1.41 Market_Val


Predictor       Coef  SE Coef      T      P
Constant        1.36    13.77   0.10  0.922
Market_Val   1.40827  0.03693  38.13  0.000


S = 68.7575   R-Sq = 95.2%   R-Sq(adj) = 95.1%


Analysis of Variance

Source          DF       SS       MS        F      P
Regression       1  6874024  6874024  1454.02  0.000
Residual Error  74   349842     4728
Total           75  7223866
```

(d) Interpret the y-intercept of the least squares line. Does it have a practical meaning for this application? Explain.
$\hat{\beta}_0 = 1.36$. The mean sale price when the appraised value is 0 is estimated to be 1.36 or \$1,360. Since $x = 0$ (appraised value = 0) is not in the observed range, this value has no meaning.

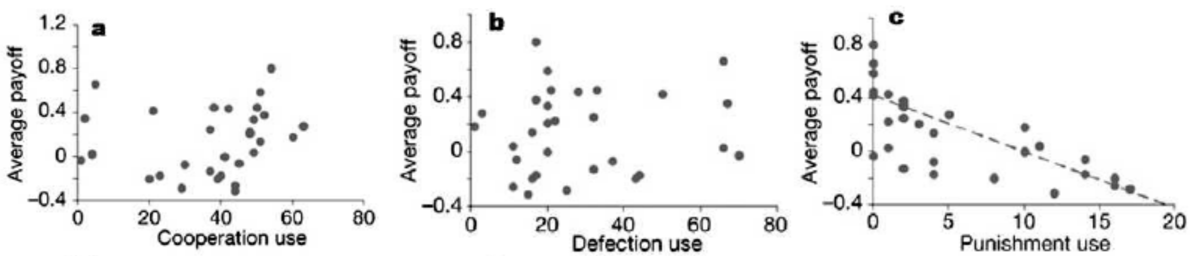
Ans: No meaning

(e) Interpret the slope of the least squares line. Over what range of x is the interpretation meaningful?
$\hat{\beta}_1 = 1.40827$. For each unit (\$1,000) increase in appraised value, the mean sale price is estimated to increase by 1.40827 (\$1,408.27).

(f) Use the least squares model to estimate the mean sale price of a property appraised at \$300,000.
\$300,000 means $x = 300$. $\hat{y} = 1.36 + 1.40827(300) = 423.841$. The estimated mean sale price for a house appraised at \$300,000 is \$423,841.

Ans: \$423,841

3.11 **In business, do nice guys finish first or last?** In baseball, there is an old saying that "nice guys finish last." Is this true in the business world? Researchers at Harvard University attempted to answer this question and reported their results in *Nature* (March 20, 2008). In the study, Boston-area college students repeatedly played a version of the game "prisoner's dilemma," where competitors choose cooperation, defection, or costly punishment. (Cooperation meant paying 1 unit for the opponent to receive 2 units; defection meant gaining 1 unit at a cost of 1 unit for the opponent; and punishment meant paying 1 unit for the opponent to lose 4 units.) At the conclusion of the games, the researchers recorded the average payoff and the number of times cooperation, defection, and punishment were used for each player. The scattergrams (p. 102) plot average payoff (y) against level of cooperation use, defection use, and punishment use, respectively.



(a) Consider cooperation use (x) as a predictor of average payoff (y). Based on the scattergram, is there evidence of a linear trend?

Ans: No

(b) Consider defection use (x) as a predictor of average payoff (y). Based on the scattergram, is there evidence on a linear trend?

Ans: No

(c) Consider punishment use (x) as a predictor of average payoff (y). Based on the scattergram, is there evidence of a linear trend?

Ans: Yes

(d) Refer to part (c). Is the slope of the line relating punishment use (x) to average payoff (y) positive or negative?

Ans: Negative

(e) The researchers conclude that "winners don't punish." Do you agree? Explain.
Yes. Winners tend to punish less than non-winners.

Ans: Yes

3.18 **Learning the mechanics**. Find SSE, s^2, and s for the least squares lines in the following exercises. Interpret the value of s.

(a) Exercise 3.6

$$SS_{yy} = \sum y^2 - n(\bar{y})^2 = 68 - 6(3)^2 = 68 - 54 = 14$$

$$SSE = SS_{yy} - \hat{\beta}_1 SS_{xy} = 14 - (.8571)(15) = 1.1435$$

$$s^2 = \frac{SSE}{n-2} = \frac{1.1435}{6-2} = .285875, s = \sqrt{.285875} = .5347$$

We expect most of the sample y-values to fall within $2s = 2(.5347) = 1.0694$ of their least squares predicted values.

Ans: SSE = 1.1435, s^2 = .285875, s = .5347

3.24 **Predicting home sales price.** Refer to the data on sale prices and total appraised values of 76 residential properties in upscale Tampa, Florida, neighborhood, Exercise 3.8 (p. 100). An SPSS simple linear regression printout for the analysis is reproduced at the bottom of the page.

SPSS Output for Exercise 3.24

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B Lower Bound	Upper Bound
1	(Constant)	1.359	13.768		.099	.922	-26.075	28.792
	Market_Val	1.408	.037	.975	38.132	.000	1.335	1.482

a. Dependent Variable: Sale_Price

(a) Use the printout to determine whether there is a positive linear relationship between appraised property value x and sale price y for residential properties sold in this neighborhood. That is, determine if there is sufficient evidence (at $\alpha = .01$) to indicate that β_1, the slope of the straight-line model, is positive.

$H_0: \beta_1 = 0, H_a: \beta_1 > 0. t = \frac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}} = \frac{1.408-0}{.037} = 38.054$. Since the observed value of the test statistic falls in the rejection region ($t = 38.132 > 38.054$), H_0 is rejected. There is sufficient evidence to indicate that y is positively linearly related to x at $\alpha = .01$.

Ans: Reject null hypothesis ($\boldsymbol{\beta_1 > 0}$)

(b) Find a 95% confidence interval for the slope, β_1, on the printout. Interpret the result practically.

The 99% confidence interval is: $\hat{\beta}_1 \pm t_{.005} s_{\hat{\beta}_1} = (1.335, 1.482)$

We are 99% confident that the increase in the mean number of sales price for each additional unit of market price ($1,000) increase is between 1.335 and 1.482. This implies that as the market price increases, the sales price also increases.

Ans: $\boldsymbol{(1.335, 1.482)}$

(c) What can be done to obtain a narrower confidence interval in part (b)?

Ans: Make $\boldsymbol{\alpha}$ < .01

3.28 **Manage therapy for boxers.** The *British Journal of Sports Medicine* (April 2000) published a study of the effect of massage on boxing performance. Two variables measured on the boxers were blood lactate concentration (mM) and the boxer's perceived recovery (28-point scale). based on information provided in the article, the data in the table were obtained for 16 five-round boxing performances, where a massage was given to the boxer between rounds. Conduct a test to determine whether blood lactate level (y) is linearly related to perceived recovery (x). Use $\alpha = .10$.

BOXING2

BLOOD LACTATE LEVEL	PERCEIVED RECOVERY
3.8	7
4.2	7
4.8	11
4.1	12
5.0	12
5.3	12
4.2	13
2.4	17
3.7	17
5.3	17
5.8	18
6.0	18
5.9	21
6.3	21
5.5	20
6.5	24

Source: Hemmings, B., Smith, M., Graydon, J., and Dyson, R. "Effects of massage on physiological restoration, perceived recovery, and repeated sports performance," *British Journal of Sports Medicine*, Vol. 34, No. 2, Apr. 2000 (data adapted from Figure 3).

$$\sum x = 78.8, \sum y = 247, \sum x^2 = 406.84, \sum y^2 = 4193$$

$$\bar{x} = 78.8 \div 16 = 4.925, \bar{y} = 247 \div 16 = 15.4375, \sum xy = 1264.6$$

$$SS_{xx} = \sum x^2 - n(\bar{x})^2 = 406.84 - 16(4.925)^2 = 406.84 - 388.09 = 18.75$$

$$SS_{yy} = \sum y^2 - n(\bar{y})^2 = 4193 - 16(15.4375)^2 = 4193 - 3813.0625 = 379.9375$$

$$SS_{xy} = \sum xy - n\bar{x}\bar{y} = 1264.6 - 16(4.925)(15.4375) = 1264.6 - 1216.475 = 48.125$$

$$\hat{\beta}_1 = \frac{SS_{xx}}{SS_{xy}} = \frac{18.75}{48.125} = .3896, \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 15.4375 - (.3896)(4.925) = 13.519$$

$$SSE = SS_{yy} - \hat{\beta}_1 SS_{xy} = 379.9375 - .3896(48.125) = 361.1875$$

$$s^2 = \frac{SSE}{n-2} = \frac{361.1875}{16-2} = 25.799, s = \sqrt{.25.799} = 5.0793, t = \frac{\hat{\beta}_1 - 0}{s_{\hat{\beta}_1}} = \frac{.3896-0}{5.0793} = .00767$$

Since the rejection region is $t = 1.761 > .00767$, H_0 is not rejected. There is sufficient evidence to indicate that perceived recovery is positively linearly related to blood lactate level at $\alpha = .10$.

Ans: Yes

3.40 **Predicting home sales price.** Refer to the data on sale prices and total appraised values of 76 residential properties recently sold in an upscale Tampa, Florida, neighborhood, Exercise 3.8 (p. 100). The MINITAB simple linear regression printout relating sale price (y) to appraised property (market) value (x) is reproduced on the next page, followed by a MINITAB correlation printout.

MINITAB Output for Exercise 3.40

Regression Analysis: Sale_Price versus Market_Val

```
The regression equation is
Sale_Price = 1.4 + 1.41 Market_Val


Predictor      Coef  SE Coef      T      P
Constant       1.36    13.77   0.10  0.922
Market_Val  1.40827  0.03693  38.13  0.000


S = 68.7575   R-Sq = 95.2%   R-Sq(adj) = 95.1%


Analysis of Variance

Source          DF       SS       MS        F      P
Regression       1  6874024  6874024  1454.02  0.000
Residual Error  74   349842     4728
Total           75  7223866
```

Correlations: Sale_Price, Market_Val

```
Pearson correlation of Sale_Price and Market_Val = 0.975
P-Value = 0.000
```

(a) Find the coefficient of correlation between appraised property value and sale price on the printout. Interpret this value.

Since $r = 0.975$ and is close to 1, there is a very strong positive linear relationship between sales price and appraised price.

Ans: 0.975

(b) Find the coefficient of determination between appraised property value and sale price on the printout. Interpret this value.

$r^2 = 95.2\%$. This means that 95.2 of the sample variance of the sale prices around the sample mean is explained by the linear relationship between sale price and appraised value.

Ans: 95.2%