

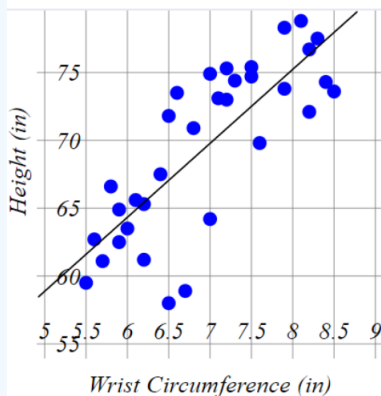
## 12.5: Chapter 12 Exercises

### ? Chapter 12 Exercises

- The correlation coefficient,  $r$ , is a number between \_\_\_\_\_.
  - 1 and 1
  - 10 and 10
  - 0 and 10
  - 0 and  $\infty$
  - 0 and 1
  - $-\infty$  and  $\infty$
- To test the significance of the correlation coefficient, we use the t-distribution with how many degrees of freedom?
  - $n - 1$
  - $n$
  - $n + 1$
  - $n - 2$
  - $n_1 + n_2 - 2$
- What are the hypotheses for testing to see if a correlation is statistically significant?
  - $H_0 : r = 0 \quad H_1 : r \neq 0$
  - $H_0 : \rho = 0 \quad H_1 : \rho \neq 0$
  - $H_0 : \rho = \pm 1 \quad H_1 : \rho \neq \pm 1$
  - $H_0 : r = \pm 1 \quad H_1 : r \neq \pm 1$
  - $H_0 : \rho = 0 \quad H_1 : \rho = 1$
- The coefficient of determination is a number between \_\_\_\_\_.
  - 1 and 1
  - 10 and 10
  - 0 and 10
  - 0 and  $\infty$
  - 0 and 1
  - $-\infty$  and  $\infty$
- Which of the following is not a valid linear regression equation?
  - $\hat{y} = -5 + \frac{2}{9}x$
  - $\hat{y} = 3x + 2$
  - $\hat{y} = \frac{2}{9} - 5x$
  - $\hat{y} = 5 + 0.4x$
- Body frame size is determined by a person's wrist circumference in relation to height. A researcher measures the wrist circumference and height of a random sample of individuals. The data is displayed below. Use  $\alpha = 0.05$ .

Regression Statistics	
Multiple R	0.7938
R Square	0.6301
Adjusted R Square	0.6182
Standard Error	3.8648
Observations	33

	Coefficients	Standard Error	t Stat	P-value
Intercept	31.6304	5.2538	6.0205	1.16E-06
Circumference	5.4496	0.7499	7.2673	3.55E-08



- What is the value of the test statistic to see if the correlation is statistically significant?
  - 6.0205
  - 1.16E-06

- iii. 3.55E-08
- iv. 5.2538
- v. 7.2673
- vi. 0.7499
- vii. 0.7938

b) What is the correct p-value and conclusion for testing if there is a significant correlation?

- i. 1.16E-06; there is a significant correlation.
- ii. 3.55E-08; there is a significant correlation.
- iii. 1.16E-06; There is not a significant correlation.
- iv. 3.55E-08, There is not a significant correlation.
- v. 0.7938, There is a significant correlation.
- vi. 0.7938, There is not a significant correlation.

7. Bone mineral density and cola consumption have been recorded for a sample of patients. Let  $x$  represent the number of colas consumed per week and  $y$  the bone mineral density in grams per cubic centimeter. Assume the data is normally distributed. Calculate the correlation coefficient.

$x$	1	2	3	4	5	6	7	8	9	10	11
$y$	0.883	0.8734	0.8898	0.8852	0.8816	0.863	0.8634	0.8648	0.8552	0.8546	0.862

8. A teacher believes that the third homework assignment is a key predictor of how well students will do on the midterm. Let  $x$  represent the third homework score and  $y$  the midterm exam score. A random sample of last term's students were selected and their grades are shown below. Assume scores are normally distributed. Use  $\alpha = 0.05$ .

H W 3	Midterm	HW3	Midterm	HW3	Midterm
1 3 . 1	59	6.4	43	20	86
2 1 . 9	87	20.2	79	15.4	73
8 . 8	53	21.8	84	25	93
2 4 . 3	95	23.1	92	9.7	52
5 . 4	39	22	87	15.1	70
1 3 . 2	66	11.4	54	15	65
2 0 . 9	89	14.9	71	16.8	77
1 8 . 5	78	18.4	76	20.1	78

- a) State the hypotheses to test for a significant correlation.
- b) Compute the correlation coefficient.
- c) Compute the p-value to see if there is a significant correlation.
- d) State the correct decision.
- e) Is there a significant correlation?
- f) Compute the coefficient of determination.
- g) Write a sentence interpreting  $R^2$ .
- h) Does doing poorly on homework 3 cause a student to do poorly on the midterm exam? Explain.
- i) Find the standard error of estimate.

9. The sum of the residuals should be \_\_\_\_\_.

- a) a
- b) b
- c) 0
- d) 1
- e)  $r$

10. An object is thrown from the top of a building. The following data measure the height of the object from the ground for a five-second period. Calculate the correlation coefficient.

Seconds	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
---------	-----	---	-----	---	-----	---	-----	---	-----	---

Height	112.5	110.875	106.8	100.275	91.3	79.875	70.083	59.83	30.65	0
--------	-------	---------	-------	---------	------	--------	--------	-------	-------	---

11. A teacher believes that the third homework assignment is a key predictor of how well students will do on the midterm. Let  $x$  represent the third homework score and  $y$  the midterm exam score. A random sample of last term's students were selected and their grades are shown below. Assume scores are normally distributed. Use  $\alpha = 0.05$ .

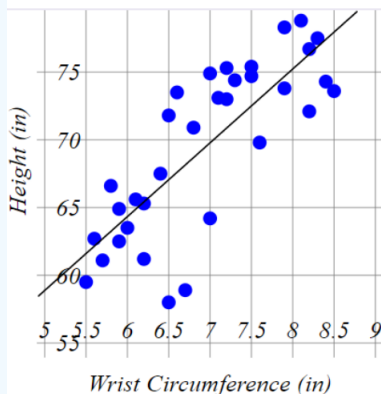
H W 3	Midterm		HW3		Midterm		HW3		Midterm
1 3 . 1	59		6.4		43		20		86
2 1 . 9	87		20.2		79		15.4		73
8 . 8	53		21.8		84		25		93
2 4 . 3	95		23.1		92		9.7		52
5 . 4	39		22		87		15.1		70
1 3 . 2	66		11.4		54		15		65
2 0 . 9	89		14.9		71		16.8		77
1 8 . 5	78		18.4		76		20.1		78

- Compute the regression equation.
- Compute the predicted midterm score when the homework 3 score is 15.
- Compute the residual for the point (15, 65).
- Find the 95% prediction interval for the midterm score when the homework 3 score is 15.

12. Body frame size is determined by a person's wrist circumference in relation to height. A researcher measures the wrist circumference and height of a random sample of individuals. The data is displayed below.

Regression Statistics	
Multiple R	0.7938
R Square	0.6301
Adjusted R Square	0.6182
Standard Error	3.8648
Observations	33

	Coefficients	Standard Error	t Stat	P-value
Intercept	31.6304	5.2538	6.0205	1.16E-06
Circumference	5.4496	0.7499	7.2673	3.55E-08



- a) Which is the correct regression equation?
- $\hat{y} = 31.6304 + 5.4496$
  - $\hat{y} = 31.6304 + 5.4496x$
  - $\hat{y} = 5.4496 + 31.6304x$
  - $\hat{y} = 31.6304 + 5.2538x$
  - $y = 31.6304 + 5.4496x$
- b) What is the predicted height (in inches) for a person with a wrist circumference of 7 inches?
- c) Which number is the standard error of estimate?
- d) Which number is the coefficient of determination?
- e) Compute the correlation coefficient.
- f) What is the correct test statistic for testing if the slope is significant  $H_1 : \beta_1 \neq 0$ ?
- g) What is the correct p-value for testing if the slope is significant  $H_1 : \beta_1 \neq 0$ ?
- h) At the 5% level of significance, is there a significant relationship between wrist circumference and height?

13. Bone mineral density and cola consumption have been recorded for a sample of patients. Let  $x$  represent the number of colas consumed per week and  $y$  the bone mineral density in grams per cubic centimeter. Assume the data is normally distributed. Calculate the coefficient of determination.

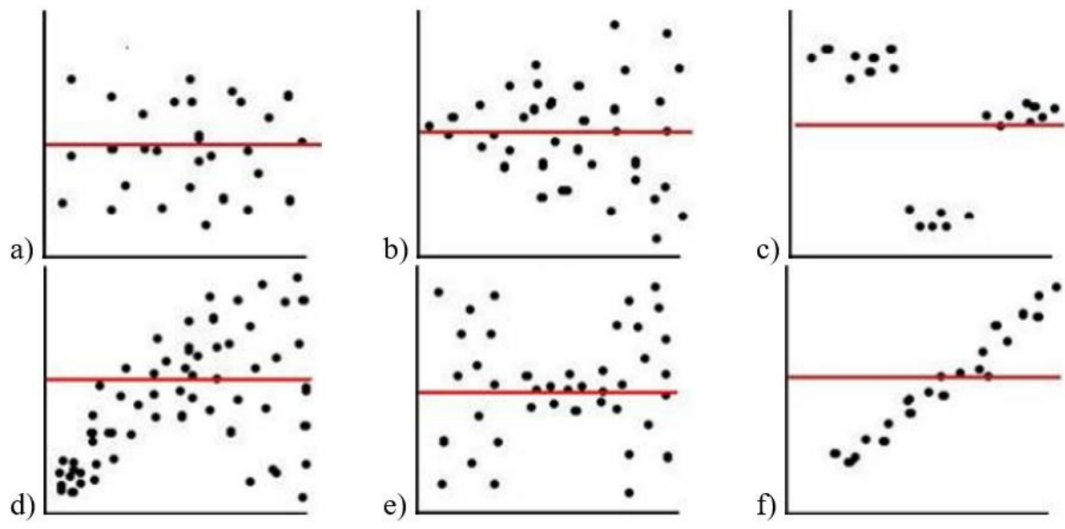
$x$	1	2	3	4	5	6	7	8	9	10	11
$y$	0.883	0.8734	0.8898	0.8852	0.8816	0.863	0.8634	0.8648	0.8552	0.8546	0.862

14. Bone mineral density and cola consumption have been recorded for a sample of patients. Let  $x$  represent the number of colas consumed per week and  $y$  the bone mineral density in grams per cubic centimeter. Assume the data is normally distributed. A regression equation for the following data is  $\hat{y} = 0.8893 - 0.0031x$ . Which is the best interpretation of the slope coefficient?

$x$	1	2	3	4	5	6	7	8	9	10	11
$y$	0.883	0.8734	0.8898	0.8852	0.8816	0.863	0.8634	0.8648	0.8552	0.8546	0.862

- For every additional average weekly soda consumption, a person's bone density increases by 0.0031 grams per cubic centimeter.
- For every additional average weekly soda consumption, a person's bone density decreases by 0.0031 grams per cubic centimeter.
- For an increase of 0.8893 in the average weekly soda consumption, a person's bone density decreases by 0.0031 grams per cubic centimeter.
- For every additional average weekly soda consumption, a person's bone density decreases by 0.8893 grams per cubic centimeter.

15. Which residual plot has the best linear regression model?

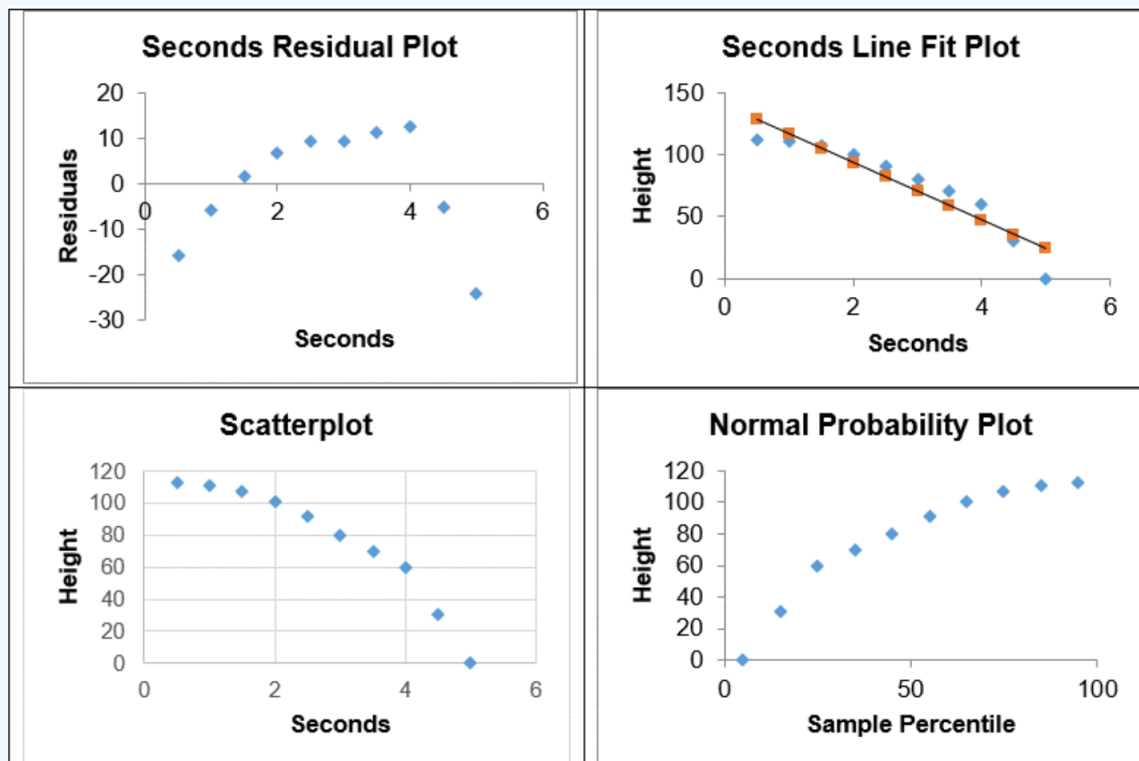


- a) a
- b) b
- c) c
- d) d
- e) e
- f) f

16. An object is thrown from the top of a building. The following data measure the height of the object from the ground for a five-second period.

Seconds	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Height	112.5	110.875	106.8	100.275	91.3	79.875	70.083	59.83	30.65	0

The following four plots were part of the regression analysis.



There is a statistically significant correlation between time and height,  $r = -0.942$ ,  $p\text{-value} = 0.0000454$ . Should linear regression be used for this data? Why or why not? Choose the correct answer.

- a) Yes, the p-value indicates that there is a significant correlation so we can use linear regression.
- b) Yes, the normal probability plot has a nice curve to it.
- c) Yes, there is a nice straight line in the line fit plot.
- d) No, there is a curve in the residual plot, normal plot and the scatterplot.

17. The following data represent the leaching rates (percent of lead extracted vs. time in minutes) for lead in solutions of magnesium chloride ( $\text{MgCl}_2$ ). Use  $\alpha = 0.05$ .

Time ( $x$ )	4	8	16	30	60	120
Percent Extracted ( $y$ )	1.2	1.6	2.3	2.8	3.6	4.4

- State the hypotheses to test for a significant correlation.
- Compute the correlation coefficient.
- Compute the p-value to see if there is a significant correlation.
- State the correct decision.
- Is there a significant correlation?
- Compute the coefficient of determination.
- Compute the regression equation.
- Compute the 95% prediction interval for 100 minutes.
- Write a sentence interpreting this interval using units and context.

18. Bone mineral density and cola consumption have been recorded for a sample of patients. Let  $x$  represent the number of colas consumed per week and  $y$  the bone mineral density in grams per cubic centimeter. Assume the data is normally distributed. What is the residual for the observed point (7, 0.8634)

$x$	1	2	3	4	5	6	7	8	9	10	11
$y$	0.883	0.8734	0.8898	0.8852	0.8816	0.863	0.8634	0.8648	0.8552	0.8546	0.862

19. A study was conducted to determine if there was a linear relationship between a person's age and their peak heart rate. Use  $\alpha = 0.05$ .

Age ( $x$ )	16	26	32	37	42	53	48	21
Peak Heart Rate ( $y$ )	220	194	193	178	172	160	174	214

- What is the estimated regression equation that relates number of hours worked and test scores for high school students.
- Interpret the slope coefficient for this problem.
- Compute and interpret the coefficient of determination.
- Compute the coefficient of nondetermination.
- Compute the standard error of estimate.
- Compute the correlation coefficient.
- Compute the 95% Prediction Interval for peak heart rate for someone who is 25 years old.

20. The following data represent the weight of a person riding a bike and the rolling distance achieved after going down a hill without pedaling.

Weight (lbs)	59	84	97	56	103	87	88	92	53	66
Rolling distance (m)	26	43	48	20	59	44	48	46	28	32

- Can it be concluded at a 0.05 level of significance that there is a linear correlation between the two variables?
- Using the regression line for this problem, find the predicted bike rolling distance for a person that weighs 110 lbs.
- Find the 99% prediction interval for bike rolling distance for a person that weighs 110 lbs.

21. Body frame size is determined by a person's wrist circumference in relation to height. A researcher measures the wrist circumference and height of a random sample of individuals. The Excel output and scatterplot are displayed below. Find the regression equation and predict the height (in inches) for a person with a wrist circumference of 7 inches. Then, compute the residual for the point (7, 75).

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	31.6304	5.2538	6.0205	1.16E-06
Circumference	5.4496	0.7499	7.2673	3.55E-08

22. It has long been thought that the length of one's femur is positively correlated to the length of one's tibia. The following are data for a classroom of students who measured each (approximately) in inches. A significant linear correlation was found between the two variables. Find the 90% prediction interval for the length of someone's tibia when it is known that their femur is 23 inches long.

Femur Length	18.7	20.5	16.2	15.0	19.0	21.3	21.0	14.3	15.8
Tibia Length	14.2	15.9	13.1	12.4	16.2	15.8	16.2	12.1	13.0

23. The data below represent the driving speed (mph) of a vehicle and the corresponding gas mileage (mpg) for several recorded instances.

Driving Speed	Gas Mileage	Driving Speed	Gas Mileage
57	21.8	62	21.5
66	20.9	66	20.5
42	25.0	67	23.0
34	26.2	52	19.4
44	24.3	49	25.3
44	26.3	48	24.3
25	26.1	41	28.4
20	27.2	38	29.6
24	23.5	26	32.5
42	22.6	24	30.8
52	19.4	21	28.8
54	23.9	19	33.5

Driving Speed	Gas Mileage	Driving Speed	Gas Mileage
60	24.8	24	25.1

- Do a hypothesis test to see if there is a significant correlation. Use  $\alpha = 0.10$ .
- Compute the standard error of estimate.
- Compute the regression equation and use it to find the predicted gas mileage when a vehicle is driving at 77 mph.
- Compute the 90% prediction interval for gas mileage when a vehicle is driving at 77 mph.

24. The following data represent the age of a car and the average monthly cost for repairs. A significant linear correlation is found between the two variables. Use the data to find a 95% prediction interval for the monthly cost of repairs for a vehicle that is 15 years old.

Age of Car (yrs)	1	2	3	4	5	6	7	8	9	10
Monthly Cost (\$)	25	34	42	45	55	71	82	88	87	90

25. In a sample of 20 football players for a college team, their weight and 40-yard-dash time in minutes were recorded.

Weight (lbs)	40-Yard Dash (min)	Weight (lbs)	40-Yard Dash (min)
285	5.95	195	4.85
185	4.99	254	5.12
165	4.92	140	4.87
188	4.77	212	5.05
160	4.52	158	4.75
156	4.67	188	4.87
256	5.22	134	4.53
169	4.95	205	4.92
210	5.06	178	4.88
165	4.83	159	4.79

- Do a hypothesis test to see if there is a significant correlation. Use  $\alpha = 0.01$ .
- Compute the standard error of estimate.
- Compute the regression equation and use it to find the predicted 40-yard-dash time for a football player that is 200 lbs.
- Compute the 99% prediction interval for a football player that is 200 lbs.
- Write a sentence interpreting the prediction interval.

26. The following data represent the enrollment at a small college during its first ten years of existence. A significant linear relationship is found between the two variables. Find a 90% prediction interval for the enrollment after the college has been open for 14 years.

Years	1	2	3	4	5	6	7	8	9	10
Enrollment	856	842	923	956	940	981	1025	996	1057	1088

27. A new fad diet called Trim-to-the-MAX is running some tests that they can use in advertisements. They sample 25 of their users and record the number of days each has been on the diet along with how much weight they have lost in pounds. The data is below. A significant linear correlation was found between the two variables. Find the 95% prediction interval for the weight lost when a person has been on the diet for 60 days.

Days on Diet	7	12	16	19	25	34	39	43	44	49
Weight Loss	5	7	12	15	20	25	24	29	33	35

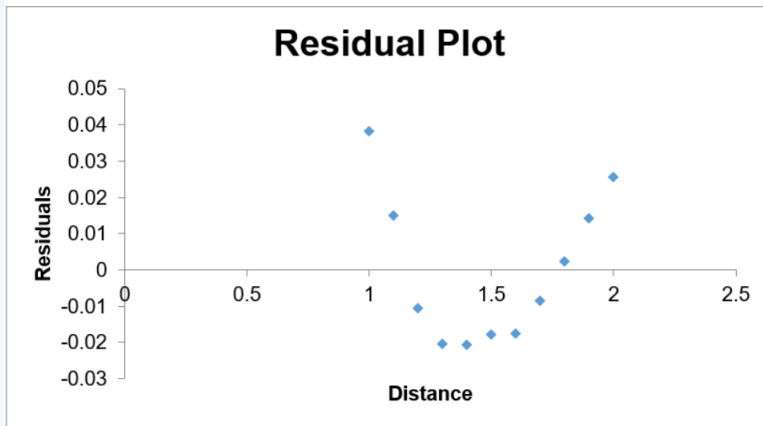
28. An elementary school uses the same system to test math skills at their school throughout the course of the 5 grades at their school. The age and score (out of 100) of several students is displayed below. A significant linear relationship is found between the student's age and their math score. Find a 90% prediction interval for the score a student would earn given that they are 5 years old.

Student Age	6	6	7	8	8	9	10	11	
Math Score	54	42	50	61	67	65	71	72	

29. The intensity (in candelas) of a 100-watt light bulb was measured by a sensing device at various distances (in meters) from the light source. A linear regression was run and the following residual plot was found.

Regression Statistics	
Multiple R	0.95936
R Square	0.920371
Adjusted R Square	0.911523
Standard Error	0.021636
Observations	11

	Coefficients	Standard Error	t Stat	P-value
Intercept	0.468618	0.031624	14.81856	1.25E-07
Distance	-0.2104	0.020629	-10.1992	3.04E-06



- Is linear regression a good model to use?
- Write a sentence explaining your answer.

30. The table below shows the percentage of adults in the United States who were married before age 24 for the years shown. A significant linear relationship was found between the two variables.

Year	1960	1965	1970	1975	1980	1985	1990	1995	2000
% Married Before 24 Years Old	52.1	51.3	45.9	46.3	40.8	38.1	34.0	32.6	28.1

- Compute the regression equation.
- Predict the percentage of adults who married before age 24 in the United States in 2015.
- Compute the 95% prediction interval for the percentage of adults who married before age 24 in the United States in 2015.

31. A nutritionist feels that what mothers eat during the months they are nursing their babies is important for healthy weight gain of their babies. She samples several of her clients and records their average daily caloric intake for the first three months of their babies' lives and also records the amount of weight the babies gained in those three months. The data are below.

Daily Calories	1523	1649	1677	1780	1852	2065	2096	2145
Baby's Weight Gain (lbs)	4.62	4.77	4.62	5.12	5.81	5.34	5.89	5.96

- Compute the regression equation.
- Test to see if the slope is significantly different from zero, using  $\alpha = 0.05$ .
- Predict the weight gain of a baby whose mother gets 2,500 calories per day.
- Compute the 95% prediction interval for the weight gain of a baby whose mother gets 2,500 calories per day.

32. The data below show the predicted average high temperature ( $^{\circ}\text{F}$ ) per month by the Farmer's Almanac in Portland, Oregon alongside the actual high temperature per month that occurred.

Farmer's Almanac	45	50	57	62	69	72	81	90	78	64	51	48
Actual High	46	52	60	61	72	78	82	95	85	68	52	49

- Compute the regression equation.
- Test to see if the slope is significantly different from zero, using  $\alpha = 0.01$ .
- Predict the high temperature in the coming year, given that the Farmer's Almanac is predicting the high to be  $58^{\circ}\text{F}$ .
- Compute the 99% prediction interval for the actual high temperature in the coming year, given that the Farmer's Almanac is predicting the high to be  $58^{\circ}\text{F}$ .

33. In a multiple linear regression problem,  $p$  represents:

- The number of dependent variables in the problem.
- The probability of success.
- The number of independent variables in the problem.
- The probability of failure.
- The population proportion.

34. The manager of a warehouse found a significant relationship exists between the number of weekly hours clocked ( $x_1$ ), the age of the employee ( $x_2$ ), and the productivity of the employee ( $y$ ) (measured in weekly orders assembled). A multiple regression was run and the following line of best fit was found:  $\hat{y} = 66.238 + 2.7048x_1 - 0.7275x_2$ . Approximate the productivity level of an employee who clocked 61 hours in a given week and is 59 years old.



35. A multiple regression test concludes that there is a linear relationship and finds the following line of best fit:  $\hat{y} = -53.247 + 12.594x_1 - 0.648x_2 + 4.677x_3$ . Use the line of best fit to approximate  $y$  when  $x_1 = 5$ ,  $x_2 = 12$ ,  $x_3 = 2$ .

36. A career counselor feels that the strongest predictors of a student's success after college are class attendance ( $x_1$ ) (recorded as a percent) and GPA ( $x_2$ ). To test this, she samples clients after they have found a job and lets the dependent variable be the number of weeks it took them to find a job. A multiple regression was run and the following line of best fit was found  $\hat{y} = 38.6609 - 0.3345x_1 - 1.0743x_2$ . Predict the number of weeks it will take a client to find a job, given that she had a 94% ( $x_1 = 94$ ) attendance rate in college and a 3.82 GPA.

37. A study conducted by the American Heart Association provided data on how age, blood pressure and smoking relate to the risk of strokes. The following data is the SPSS output with Age in years ( $x_1$ ), Blood Pressure in mmHg ( $x_2$ ), Smoker (0 = Nonsmoker, 1 = Smoker) ( $x_3$ ) and the Risk of a Stroke as a percent ( $y$ ).

#### SUMMARY OUTPUT

Regression Statistics	
Multiple R	
R Square	0.873
Adjusted R Square	
Standard Error	5.75657
Observations	20

#### ANOVA

	df	SS	MS	F	Significance F
Regression	3	3660.740	120.247	36.823	2.04E-07
Residual	16	530.210	33.138		
Total	19	4190.950			

	Coefficients	Standard Error	t Stat	P-value
Intercept	-91.759	15.223	-6.028	1.7568E-05
Age	1.077	0.166	6.488	7.4848E-07
Blood Pressure	0.252	0.045	5.568	4.2433E-05
Smoker	8.740	3.001	2.912	0.0102

- Use  $\alpha = 0.05$  to test the claim that the regression model is significant. State the hypotheses, test statistic, p-value, decision and summary.
- Use the estimated regression equation to predict the stroke risk for a 70-year-old smoker with a blood pressure of 180.
- Interpret the slope coefficient for age.
- Find the adjusted coefficient of determination.

38. Use technology to run a multiple linear regression with a dependent variable of annual salary for adults with full-time jobs in San Francisco and independent variables of years of education completed and age. The data for the sample used is below.

Annual Salary	Years of Education	Age
82,640	17	34
95,854	18	42
152,320	21	39
49,165	13	25
31,120	12	31
67,500	16	32
42,590	12	28
57,245	14	55
58,940	16	45
67,250	18	40
56,120	16	39
38,955	12	34
74,650	16	33
53,495	16	29
67,210	16	30
79,365	16	50
96,045	18	51
78,472	14	60
124,975	21	52
43,125	12	36

- Test to see if the overall model is significant using  $\alpha = 0.05$ .
- Compute the predicted salary for a 35-year-old with 16 years of education.
- Compute the adjusted coefficient of determination.
- Are all the independent variables statistically significant? Explain.

39. A study was conducted to determine if there was a linear relationship between a person's weight in pounds with their gender, height and activity level. A person's gender was recorded as a 0 for anyone who identified as male and 1 for those who did not identify as male. Height was measured in inches. Activity level was coded as 1, 2, or 3; the more active the person was, the higher the value.

### SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	
R Square	0.667
Adjusted R Square	
Standard Error	13.9353
Observations	92

### ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	34195	11398.33	58.6958	6.2026E-21
Residual	88	17089	194.1932		
Total	91	51284			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-81.5	43.96	-1.854	0.0671
Gender	-15.1	4.286	-3.5231	0.0007
Height	3.7	0.5704	6.4867	4.97E-09
Activity Level	-3.08	2.443	-1.2607	0.2107

- Interpret the slope coefficient for height.
- Predict the weight for a male who is 70 inches tall and has an activity level of 2.
- Calculate the adjusted coefficient of determination.

40. A professor claims that the number of hours a student studies for their final paired with the number of homework assignments completed throughout the semester (out of 20 total assignment opportunities) is a good predictor of a student's final exam grade. She collects a sample of students and records the following data.

Final Exam Score	# Hours Studied	HW Completed
85	5.2	19
74	5.1	18
79	4.9	16
62	3.7	12
96	9	17
52	1	12
73	2	15
81	2	16
90	7.25	18
79	4.5	14
83	3	20
76	3.2	17
92	6.8	20
84	6.2	18
50	4.1	10

- Test the professor's claim using  $\alpha = 0.05$ .
- Compute the multiple correlation coefficient.
- Interpret the slope coefficient for number of hours studied.
- Compute the adjusted coefficient of determination.

### Solutions to Odd-Numbered Exercises

- a
- b
- d
- $r = 0.8241$
- c
- a)  $\hat{y} = 25.6472 + 2.8212x$   
 b) 67.9657  
 c) -2.9657  
 d)  $61.3248 < y < 74.6065$

13.  $R^2 = 0.679$

15. a

17. a)  $H_0 : \rho = 0; H_1 : \rho \neq 0$

b)  $r = 0.9403$

c)  $0.0052$

d) Reject  $H_0$

e) Yes

f)  $R^2 = 0.8842$

g)  $\hat{y} = 1.6307 - 0.0257x$

h)  $2.6154 < y < 5.7852$

i) We can be 95% confident that the predicted percent of lead extracted in solutions of magnesium chloride at 100 minutes is anywhere between 2.6154 and 5.7852.

19. a)  $\hat{y} = 241.8127 - 1.5618x$

b) Every year a person ages, their peak heart rate decreases by an average of 1.5618.

c)  $R^2 = 0.93722$

d)  $0.06278$

e)  $5.6924$

f)  $r = -0.9681$

21.  $5.2224$

23. a)  $H_0 : \rho = 0; H_1 : \rho \neq 0; t = -5.2514; \text{p-value} = 0.000022$

Reject  $H_0$ . There is a significant correlation between driving speed of a vehicle and the corresponding gas mileage.

b)  $2.49433$

c)  $\hat{y} = 32.40313 - 0.1662x$

d)  $14.8697 < y < 24.3424$

25. a)  $H_0 : \rho = 0; H_1 : \rho \neq 0; t = 6.9083; \text{p-value} = 0.000007$

Reject  $H_0$ . There is a significant correlation between a football player's weight and 40-yard-dash time.

b)  $s = 0.160595$

c)  $\hat{y} = 3.722331 - 0.006396x; 5.0016$

d)  $4.527 < y < 5.476$

e) We can be 99% confident that the predicted time for all 200-pound football player's running time for the 40-yard-dash is between 4.527 and 5.476 minutes.

27.  $36.605 < y < 47.740$

29. a) No

b) The p-value = 0.00000304 suggests that there is a significant linear relationship between intensity (in candelas) of a 100-watt light bulb was measured by a sensing device at various distances (in meters) from the light source. However, the residual plot clearly shows a nonlinear relationship. Even though we can fit a straight line through the points, we would get a better fit with a curve.

31. a)  $\hat{y} = 1.76267 - 0.001883x$

b)  $H_0 : \beta_1 = 0; H_1 : \beta_1 \neq 0; t = 5.0668; \text{p-value} = 0.0015$

Reject  $H_0$ . There is significant linear relationship between a nursing baby's weight gain and the calorie intake of the mother.

c)  $6.4693 \text{ lbs}$

d)  $5.571 < y < 7.368$

33. c

35.  $11.301$

37. a)  $H_0 : \beta_1 = \beta_2 = \beta_3 = 0; H_1 : \text{At least one slope is not zero}; F = 36.823; \text{p-value} = 0.000000204$

Reject  $H_0$ . There is a significant relationship between a person's age, blood pressure and smoking status and the risk of a stroke.

b)  $37.731$

c) For each year a person ages, their chance of a stroke increases by 1.077%.

d)  $84.9$

39. a) For each additional inch in height, the predicted weight would increase 3.7 pounds.

b)  $171.34 \text{ lbs}$

c)  $65.56$