

E Chapter 12 Exercises

Answer each question.

1. The following table gives the heights and corresponding batting averages for a sample of 12 professional baseball players.

Heights and Batting Averages												
Height (in Inches), x	70	73	73	71	68	76	71	75	73	73	75	73
Batting Average, y	.248	.273	.245	.202	.304	.349	.226	.250	.233	.273	.237	.229

- a. Calculate the correlation coefficient.
 - b. Is the correlation coefficient statistically significant at the 0.05 level? At the 0.01 level?
 - c. Should a regression line be used to model the data and make predictions? Why or why not?
2. The following table gives the college GPAs and corresponding annual incomes after graduation for a sample of nine young adults.

College GPAs and Annual Incomes after Graduation									
GPA, x	3.1	3.8	2.9	1.8	2.7	3.5	3.2	2.2	4.0
Annual Income, y	\$38,000	\$36,500	\$39,500	\$29,000	\$33,000	\$40,500	\$26,000	\$30,000	\$61,000

- a. Before calculating the correlation coefficient, what kind of relationship would you expect to see between the two variables?
 - b. Calculate the correlation coefficient and coefficient of determination.
 - c. Is r statistically significant at the 0.05 level? At the 0.01 level?
 - d. Suppose that your college GPA is 3.3. Using the regression model for these data, make a prediction for your future annual income.
3. The interest rate that day for a 30-year mortgage and the corresponding number of new loans taken out at a large regional bank are given for 12 consecutive days.

Interest Rates and Numbers of New Loans Taken Out												
Interest Rate (%), x	4.54	4.57	4.62	4.62	4.59	4.63	4.68	4.75	4.83	4.90	5.01	5.01
New Loans, y	10	8	7	7	7	9	9	8	6	5	4	4

- a. Before calculating the correlation coefficient, what kind of relationship would you expect to see between the two variables?
- b. Calculate the correlation coefficient and coefficient of determination.
- c. Is r statistically significant at the 0.05 level? At the 0.01 level?
- d. Suppose that the average interest rate for a given month is 4.75%. Using the regression model for these data, make a prediction for the number of new loans taken out.

4. The following table gives the average price of gas and the corresponding SUV sales at a local car dealership for ten randomly selected months in 2003.

Gas Prices and SUV Sales										
Price of Gas (per Gallon), x	\$1.26	\$1.24	\$1.51	\$1.37	\$1.53	\$1.43	\$1.35	\$1.49	\$1.42	\$1.58
SUVs Sold, y	45	44	36	45	40	43	47	39	47	37

- Before calculating the correlation coefficient, what kind of relationship would you expect to see between the two variables?
 - Calculate the correlation coefficient and coefficient of determination.
 - Is r statistically significant at the 0.05 level? At the 0.01 level?
 - Suppose that the average gas price for one month is \$1.40. How many SUVs should the car dealership expect to sell?
 - Obviously, these data from 2003 should not be used for predictions this year. However, can you think of ways that the data might be useful for today's car dealerships?
5. A lifeguard collected the following data while the lifeguard was observing the beach on the first day of each month, over the course of a year.

Numbers of Swimsuits and Seagulls												
Swimsuits, x	0	8	11	28	77	120	150	250	100	44	12	5
Seagulls, y	2	78	92	106	102	200	300	376	298	55	15	3

- From analyzing the data above, can you conclude that there is a significant linear relationship between the two variables at the 0.05 level? At the 0.01 level?
 - Could you predict how many seagulls would be visible if you can see 200 swimsuits? What about 500 swimsuits?
 - What reasonable conclusions can you make regarding your findings?
6. Consider the following data from a recent study that included 14 participants who have used a cell phone consistently over the course of several years. For each participant, the number of years they had been using a cell phone was recorded, along with an estimate of the probability that they would get cancer in their lifetime, based on a complex system of multiple factors.

Cancer Risk and Number of Years Using a Cell Phone														
Years Using a Cell Phone, x	2	3.1	1.5	0.9	4	1.5	3.9	2	5	4.4	2.1	4.9	8	6.2
Cancer Risk, y	0.35	0.11	0.45	0.28	0.55	0.78	0.55	0.38	0.74	0.81	0.15	0.09	0.66	0.2

- From analyzing the data, can you conclude that there is a significant linear relationship between the two variables at the 0.05 level? At the 0.01 level?
- Is it reasonable to make predictions of cancer risk based on cell phone usage?
- If you were a cell phone company, what would you do with the results of this study?

7. The following data were collected from a used car dealership regarding the selling prices of used cars and the numbers of miles on the cars when they were sold.

Mileages and Selling Prices of Used Cars	
Mileage, x	Selling Price, y
23,547	\$21,803
48,710	\$17,995
33,952	\$18,995
14,876	\$20,863
133,691	\$4970
30,417	\$28,500
29,104	\$22,597
54,000	\$17,600
63,079	\$14,999
133,172	\$10,995
110,000	\$8000
63,263	\$15,995

- Can you conclude that there is a significant linear relationship between the two variables the 0.05 level? At the 0.01 level?
 - Predict the selling price of a vehicle that has 42,000 miles on it.
 - How much of the variation in the selling prices is associated with the variation in the mileages on the cars when they were sold?
8. Consider the following data collected on airline tickets. The following table lists the distance from Dallas, Texas to the destination and the corresponding price of the airline ticket.

Distances and Ticket Prices to Destinations from Dallas, TX		
Distance (in Miles), x	Destination	Ticket Price, y
1573.61	New York	\$212
1329.96	Washington, DC	\$234
1316.38	Miami	\$275
782.50	Atlanta	\$231
2203.26	Seattle	\$312
967.73	Chicago	\$185
1368.18	Baltimore	\$267
453.18	Memphis	\$229
521.11	New Orleans	\$345

- Is there a significant linear relationship at the 0.05 level between distance traveled and price of a ticket from Dallas, Texas?
- Predict the price of a ticket to San Francisco from Dallas assuming that there is a distance of 1465 miles between the two cities.

9. The following data represent the thread counts of various bedsheets and their corresponding prices.

Thread Counts and Prices of Bedsheets								
Thread Count, x	150	200	225	250	275	300	350	400
Price (in Dollars), y	18	21	25	28	30	31	35	45

- Calculate the sum of squared errors, SSE.
 - Calculate the standard error of estimate, S_e .
 - Construct a 95% prediction interval for the price of 350-thread-count sheets.
 - Construct a 95% confidence interval for the y -intercept of the regression line.
 - Construct a 95% confidence interval for the slope of the regression line.
10. The following data were collected on the numbers of years of post-high-school education and the annual incomes of eight people ten years after graduation from high school.

Post-High-School Education and Annual Income								
Education after High School (in Years), x	2	3	3	4	5	6	6	8
Annual Income (in Thousands of Dollars), y	35	41	38	45	65	89	95	115

- Calculate the sum of squared errors, SSE.
 - Calculate the standard error of estimate, S_e .
 - Construct a 95% prediction interval for the annual income of someone ten years after graduation from high school who had seven years of post-high-school education.
 - Construct a 95% confidence interval for the y -intercept of the regression line.
 - Construct a 95% confidence interval for the slope of the regression line.
11. In addition to the thread count of bed sheets, let's explore if the number of colors used is linearly related to the price of the sheets. Determine if a statistically significant linear relationship exists between these explanatory variables and the response variable at the 0.05 level. If the linear relationship is significant, write the multiple regression equation that best fits the data.

Thread Counts, Colors, and Prices of Bed Sheets		
Thread Count, x_1	Number of Colors, x_2	Price (in Dollars), y
150	2	18
200	1	21
225	2	25
250	4	28
275	3	30
300	2	31
350	1	35
400	4	45

12. In addition to the number of years of post-high-school education, the number of years of work experience also influences a person's salary. Determine if a statistically significant linear relationship exists between these explanatory variables and the response variable at the 0.05 level. If the linear relationship is significant, write the multiple regression equation that best fits the data.

Education, Work Experience, and Annual Income		
Education after High School (in Years), x_1	Work Experience (in Years), x_2	Annual Income (in Thousands of Dollars), y
2	8	35
3	7	41
3	5	38
4	6	45
5	6	65
6	5	89
6	4	95
8	4	115