

39. An Environmental Protection Agency study of 12 automobiles revealed a correlation of 0.47 between engine size and emissions. At the .01 significance level, can we conclude that there is a positive association between these variables? What is the p -value? Interpret.
40. A suburban hotel derives its gross income from its hotel and restaurant operations. The owners are interested in the relationship between the number of rooms occupied on a nightly basis and the revenue per day in the restaurant. Below is a sample of 25 days (Monday through Thursday) from last year showing the restaurant income and number of rooms occupied.

Day	Income	Occupied	Day	Income	Occupied
1	\$1,452	23	14	\$1,425	27
2	1,361	47	15	1,445	34
3	1,426	21	16	1,439	15
4	1,470	39	17	1,348	19
5	1,456	37	18	1,450	38
6	1,430	29	19	1,431	44
7	1,354	23	20	1,446	47
8	1,442	44	21	1,485	43
9	1,394	45	22	1,405	38
10	1,459	16	23	1,461	51
11	1,399	30	24	1,490	61
12	1,458	42	25	1,426	39
13	1,537	54			

Use a statistical software package to answer the following questions.

- Does the breakfast revenue seem to increase as the number of occupied rooms increases? Draw a scatter diagram to support your conclusion.
 - Determine the correlation coefficient between the two variables. Interpret the value.
 - Is it reasonable to conclude that there is a positive relationship between revenue and occupied rooms? Use the .10 significance level.
 - What percent of the variation in revenue in the restaurant is accounted for by the number of rooms occupied?
41. The table below shows the number of cars (in millions) sold in the United States for various years and the percent of those cars manufactured by GM.

Year	Cars Sold (millions)	Percent GM	Year	Cars Sold (millions)	Percent GM
1950	6.0	50.2	1985	15.4	40.1
1955	7.8	50.4	1990	13.5	36.0
1960	7.3	44.0	1995	15.5	31.7
1965	10.3	49.9	2000	17.4	28.6
1970	10.1	39.5	2005	16.9	26.9
1975	10.8	43.1	2010	11.6	19.1
1980	11.5	44.0			

Use a statistical software package to answer the following questions.

- Is the number of cars sold directly or indirectly related to GM's percentage of the market? Draw a scatter diagram to show your conclusion.
 - Determine the correlation coefficient between the two variables. Interpret the value.
 - Is it reasonable to conclude that there is a negative association between the two variables? Use the .01 significance level.
 - How much of the variation in GM's market share is accounted for by the variation in cars sold?
42. For a sample of 32 large U.S. cities, the correlation between the mean number of square feet per office worker and the mean monthly rental rate in the central business district is -0.363 . At the .05 significance level, can we conclude that there is a negative association in the population between the two variables?

43. The following data from the 2010 NFL football season report the number of points scored and points allowed for each of the 32 NFL teams. 

Team	Points Scored	Points Allowed	Team	Points Scored	Points Allowed
Arizona	289	434	Miami	273	333
Atlanta	414	288	Minnesota	281	348
Baltimore	357	270	New England	518	313
Buffalo	283	425	New Orleans	384	307
Carolina	196	408	NY Giants	394	347
Chicago	334	286	NY Jets	367	304
Cincinnati	322	395	Oakland	410	371
Cleveland	271	332	Philadelphia	439	377
Dallas	394	436	Pittsburgh	375	232
Denver	344	471	San Diego	441	322
Detroit	362	369	San Francisco	305	346
Green Bay	388	240	Seattle	310	407
Houston	390	427	St. Louis	289	328
Indianapolis	435	388	Tampa Bay	341	318
Jacksonville	353	419	Tennessee	356	339
Kansas City	366	326	Washington	302	377

You will want to use statistical software to perform the calculations. Assume that these are sample data.

- Determine the correlation coefficient. Are you surprised at the negative association between the variables? Interpret the relationship between "points scored" and "points allowed."
- Determine the coefficient of determination. What does the coefficient of determination say about the relationship?
- Can we conclude that there is a negative association between "points scored" and "points allowed"? Use the .05 significance level.

44.  Meryl's Apparel is an upscale chain of women's clothing stores, located primarily in the southwest United States. Due to recent success, Meryl's top management is planning to expand by locating new stores in other regions of the country. The director of planning has been asked to study the relationship between yearly sales and the store size. As part of the study, the director selects a sample of 25 stores and determines the size of the store in square feet and the sales for last year. The sample data follow. The use of statistical software is suggested. 

Store Size (thousands of square feet)	Sales (millions \$)	Store Size (thousands of square feet)	Sales (millions \$)
3.7	9.18	0.4	0.55
2.0	4.58	4.2	7.56
5.0	8.22	3.1	2.23
0.7	1.45	2.6	4.49
2.6	6.51	5.2	9.90
2.9	2.82	3.3	8.93
5.2	10.45	3.2	7.60
5.9	9.94	4.9	3.71
3.0	4.43	5.5	5.47
2.4	4.75	2.9	8.22
2.4	7.30	2.2	7.17
0.5	3.33	2.3	4.35
5.0	6.76		

- a. Draw a scatter diagram. Use store size as the independent variable. Does there appear to be a relationship between the two variables. Is it positive or negative?
- b. Determine the correlation coefficient and the coefficient of determination. Is the relationship strong or weak? Why?
- c. At the .05 significance level, can we conclude there is a significant positive correlation?
45. The manufacturer of Cardio Glide exercise equipment wants to study the relationship between the number of months since the glide was purchased and the time, in hours, the equipment was used last week. df

Person	Months Owned	Hours Exercised	Person	Months Owned	Hours Exercised
Rupple	12	4	Massa	2	8
Hall	2	10	Sass	8	3
Bennett	6	8	Karl	4	8
Longnecker	9	5	Malrooney	10	2
Phillips	7	5*	Veights	5	5

- a. Plot the information on a scatter diagram. Let hours of exercise be the dependent variable. Comment on the graph.
- b. Determine the correlation coefficient. Interpret.
- c. At the .01 significance level, can we conclude that there is a negative association between the variables?
46. The following regression equation was computed from a sample of 20 observations:

$$\hat{Y} = 15 - 5X$$

SSE was found to be 100 and SS total 400.

- a. Determine the standard error of estimate.
- b. Determine the coefficient of determination.
- c. Determine the correlation coefficient. (Caution: Watch the sign!)
47. City planners believe that larger cities are populated by older residents. To investigate the relationship, data on population and median age in 10 large cities were collected. df

City	Population (in millions)	Median age
Chicago, IL	2.833	31.5
Dallas, TX	1.233	30.5
Houston, TX	2.144	30.9
Los Angeles, CA	3.849	31.6
New York, NY	8.214	34.2
Philadelphia, PA	1.448	34.2
Phoenix, AZ	1.513	30.7
San Antonio, TX	1.297	31.7
San Diego, CA	1.257	32.5
San Jose, CA	0.930	32.6

- a. Plot this data on a scatter diagram with median age as the dependent variable.
- b. Find the correlation coefficient.
- c. A regression analysis was performed and the resulting regression equation is Median age = 31.4 + 0.272 population. Interpret the meaning of the slope.
- d. Estimate the median age for a city of 2.5 million people.
- e. Here is a portion of the regression software output. What does it tell you?

Predictor	Coef	SE Coef	T	P
Constant	31.3672	0.6158	50.94	0.000
Population	0.2722	0.1901	1.43	0.190

- f. Using the .10 significance level, test the significance of the slope. Interpret the result. Is there a significant relationship between the two variables?