8.2: Linear Correlation

Learning Objectives

- Discuss linear correlation
- Motivate and develop Pearson's correlation coefficient r
- Calculate, by hand and using technology, and interpret r
- Emphasize the need for visualization in conjunction with summary statistics

Section 8.2 Excel File (contains all of the data sets for this section)

Review and Preview

When we study bivariate quantitative data, we attempt to identify if there is a relationship between the two variables. If there is a relationship, which we call an association, we would like to describe it as well as we can.

One way to do this is to describe what happens as one of the variables changes. If as one variable increases, the other increases, we say that the variables are positively correlated. If as one variable increases, the other decreases, we say that the variables are negatively correlated. Some associations are not correlated, meaning, as one variable increases, the other variable may increase or decrease depending on which values we are considering. When we have a correlated association, we can further describe the association by determining whether or not the rate of change is constant. In other words, we could determine if a linear function would be a good model for the data. A linear function models the association well when the scatter plot forms a fairly straight path through the coordinate plane. A third way to describe the association is to assess its strength or how densely packed the points are along the path in the scatter plot.

In the previous section, we noted that visual assessments of these considerations can be influenced by the units on the data and the scale of the scatter plots. In this section, we develop an analytical tool to help us, but for best results, we should use both the visualizations and analytical tools in conjunction. As indicated in the previous section, we will restrict our discussion to identifying when an association is well-modeled by a linear function and then measuring the strength of that association. Since linear functions have a constant slope, linear associations are necessarily correlated. As such, we begin our exposition of linear correlation.

Linear Correlation

We are about to embark on the development of Pearson's correlation coefficient, r, for sample data. There is an analogous measure for population data which we will not address in this course. The goal of the correlation coefficient is to assess the strength of the correlation between two variables which are thought to have a linear association. Such a measure would be expected to be independent of both the units describing the data and the ordering of the variables. Returning to our example regarding the ages of the bride and groom on their wedding day, our computation of r should not depend on whether the data is presented in months, years, or decades, and it should not care if our variables are paired as (age of bride, age of groom) or (age of groom, age of bride).

Let us begin with the task of trying to ensure that the measure does not depend on the particular units used to describe the data. Recall that we can use the z-score to compare values within and between different sets of data because the z-score represents how many standard deviations above the mean an observation is. We defined the z-score within the context of population data, but the same concepts apply when we are studying sample data. We are using sample statistics in calculation rather than population parameters. Within this context we have the following.

$$z = \frac{x - \bar{x}}{s}$$

An astute reader will acknowledge a possible issue here. We are now dealing with bivariate data; we have not discussed the idea of means and standard deviations within this context. Luckily, this does not pose any significant issue. We can study each of the individual variables as we have done previously. If we have bivariate data with a first variable x and a second variable y, we can compute the mean and standard deviation of the variable x and then the mean and standard deviation of the variable y in order to compute the associated z-score for each component of each observation. We consider two z-score computations, one for each variable as follows.

$$z_x = \frac{x - \bar{x}}{s_x} \qquad z_y = \frac{y - \bar{y}}{s_y}$$

Text Exercise 8.2.1

1. To begin our exploration, we will consider the following bivariate data set consisting of 10 observations with variables x and y. Construct a scatter plot of bivariate data to confirm that a linear function seems appropriate as a model for the data.

Table 8.2.1: Table of values for variables x and y

x	y
4	100
8	91
12	64
18	46
20	28
22	10
28	10
32	−17
36	−62
40	−80

Answer

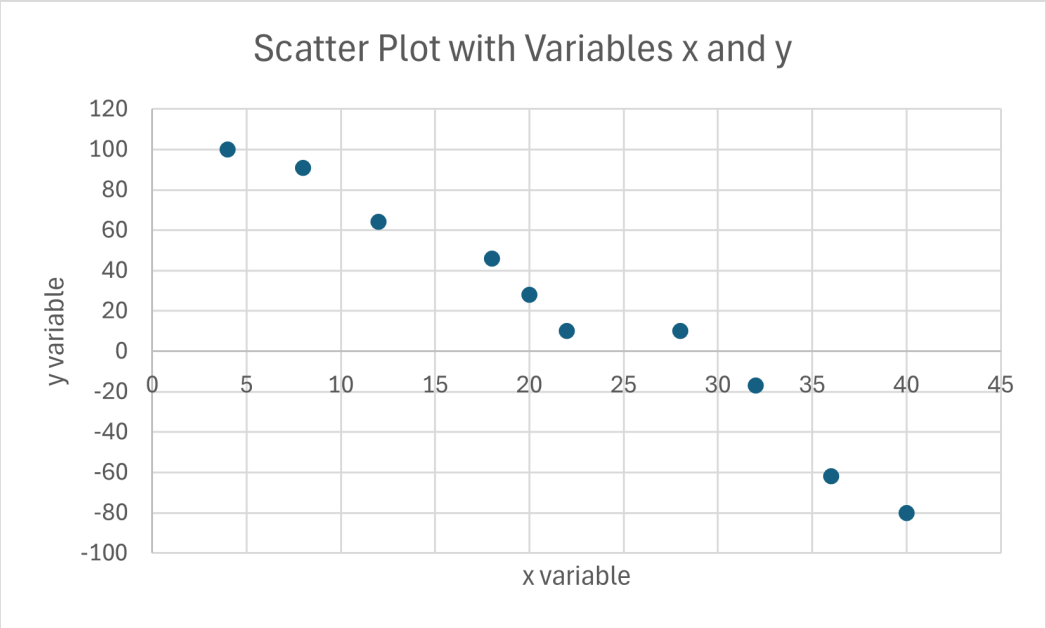


Figure 8.2.1: Scatter plot of variables x and y

After constructing a scatter plot using the variables x and y, we come to the conclusion that it is reasonable to understand the data using a linear function to model the data. It appears that the two variables are negatively correlated.

2. Compute the appropriate z-score for each component of each observation by first computing $\bar{x}$, $\bar{y}$, s_x, and s_y, and then computing the z-score according to which variable the data belongs. Use the following table as a template.

Table 8.2.2: Table of values for variables x and y

x	y	z_x	z_y

x	y	z_x	z_y
4	100		
8	91		
12	64		
18	46		
20	28		
22	10		
28	10		
32	-17		
36	-62		
40	-80		

Answer

In order to compute the appropriate means and standard deviations, we treat each column of our data set as its own separate set of data. We compute $\bar{x}$ by adding all of the values in the x column and dividing by 10 and compute $\bar{y}$ by adding all of the values in the y column and then dividing by 10. We proceed similarly for the computations of the standard deviations by treating the data from each variable separately. Using technology can make this process much quicker. We provide the appropriate values now: $\bar{x} = 22$, $s_x = 12$, $\bar{y} = 19$, and $s_y = 60$. We compute the number of standard deviations each measured value of each variable is away from the mean its particular variable.

Table 8.2.3 Table of values and calculations for variables x and y

x	y	z_x	z_y
4	100	$\frac{4-22}{12} = -\frac{3}{2}$	$\frac{100-19}{60} = \frac{27}{20}$
8	91	$\frac{8-22}{12} = -\frac{7}{6}$	$\frac{91-19}{60} = \frac{6}{5}$
12	64	$\frac{12-22}{12} = -\frac{5}{6}$	$\frac{64-19}{60} = \frac{3}{4}$
18	46	$\frac{18-22}{12} = -\frac{1}{3}$	$\frac{46-19}{60} = \frac{9}{20}$
20	28	$\frac{20-22}{12} = -\frac{1}{6}$	$\frac{18-19}{60} = \frac{3}{20}$
22	10	$\frac{22-22}{12} = 0$	$\frac{10-19}{60} = -\frac{3}{20}$
28	10	$\frac{28-22}{12} = \frac{1}{2}$	$\frac{10-19}{60} = -\frac{3}{20}$
32	-17	$\frac{32-22}{12} = \frac{5}{6}$	$\frac{-17-19}{60} = -\frac{3}{5}$
36	-62	$\frac{36-22}{12} = \frac{7}{6}$	$\frac{-62-19}{60} = -\frac{27}{20}$
40	-80	$\frac{40-22}{12} = \frac{3}{2}$	$\frac{-80-19}{60} = -\frac{33}{20}$

3. Construct a scatter plot of the z-score data and compare it with the scatter plot constructed in the first part of this text exercise.

Answer

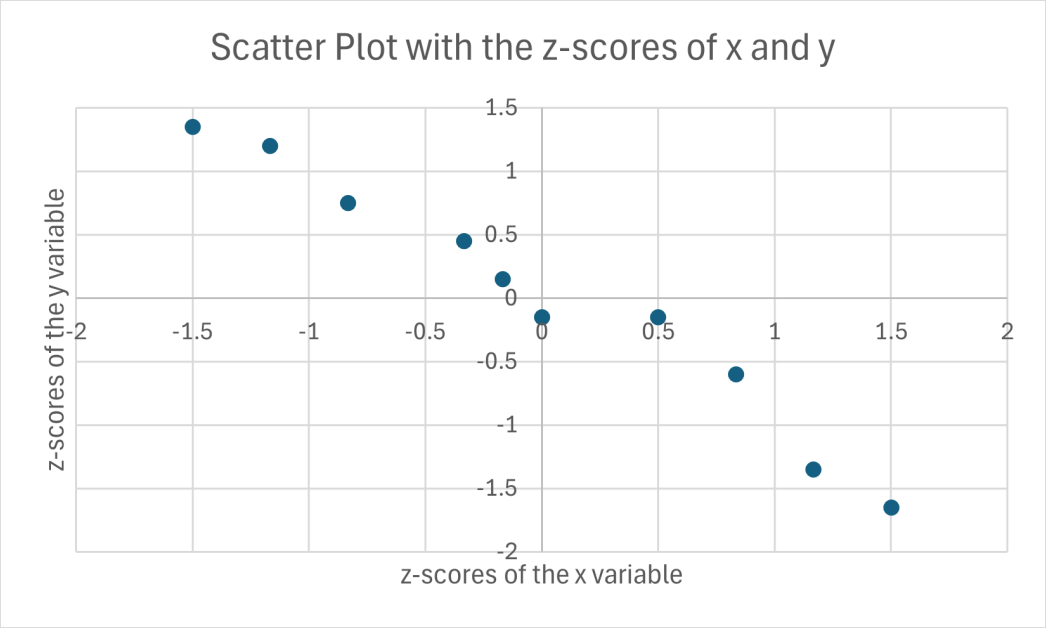


Figure 8.2.2 Scatter plot of z-scores of variables x and y

A visual analysis comparing the scatter plot using the variables z_x and z_y with the scatter plot using the variables x and $y$yields that the shape of the path looks extraordinarily similar. The relative positions of the coordinates seem to match perfectly (which they do). In considering the z-scores of our data, we have successfully removed the impact that different units would make while preserving the relationship in the data.

Let us continue to analyze the scatter plot of the variables z_x and z_y. We note that almost all of the coordinate pairs fall into 2 of the 4 quadrants of the coordinate plane. The top left quadrant, quadrant II, houses 5 of the coordinate pairs; while, the bottom right quadrant, quadrant IV, houses 4 of the coordinate pairs. The last coordinate pair falls on the boundary between quadrants III and IV. We notice that when we have a fairly strong negative correlation the majority of points land in quadrants II and IV.

Text Exercise 8.2.2

1. Recall that the bivariate data that relayed the ages of bride and groom on their wedding day displayed a positive linear correlation when we constructed the scatter plot. Using Excel to transform the data using the z-score, just as we did in the last text exercise, and then construct a scatter plot. Analyze the scatter plot by considering the points in the various quadrants. Does a similar pattern appear?

Answer

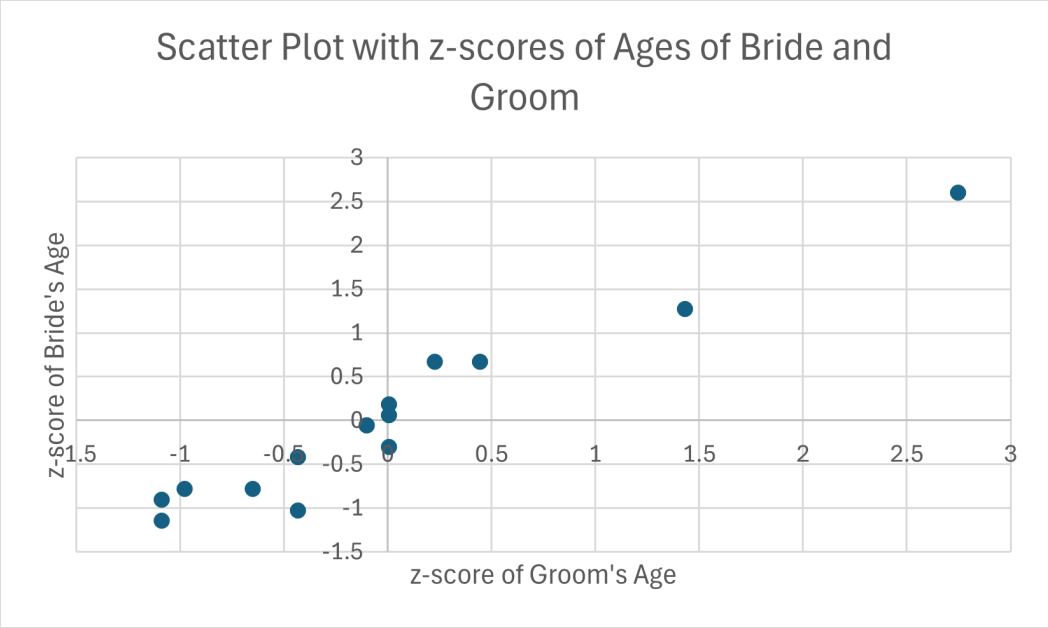


Figure 8.2.3 Scatter plot of z-scores of ages of bride and groom

The scatter plot shows that most of the points fall in quadrant I and quadrant III. There is one point that falls in quadrant IV, but it is very close to the boundary between quadrants III and IV. We see a similar trend to the scatter plot in the previous text exercise that the majority of points fall in two quadrants, but the quadrants are different. Rather than the quadrants labeled with even numbers, we have the quadrants labeled with odd numbers.

2. Now consider the scatter plot of the transformed variables from text exercise 8.1.3.1 presented below. We described the relationship between these variables as having a weaker positive correlation because there was a straight upward path through the data, but the points were not densely packed together. What do you notice about the quadrant analysis in this case?

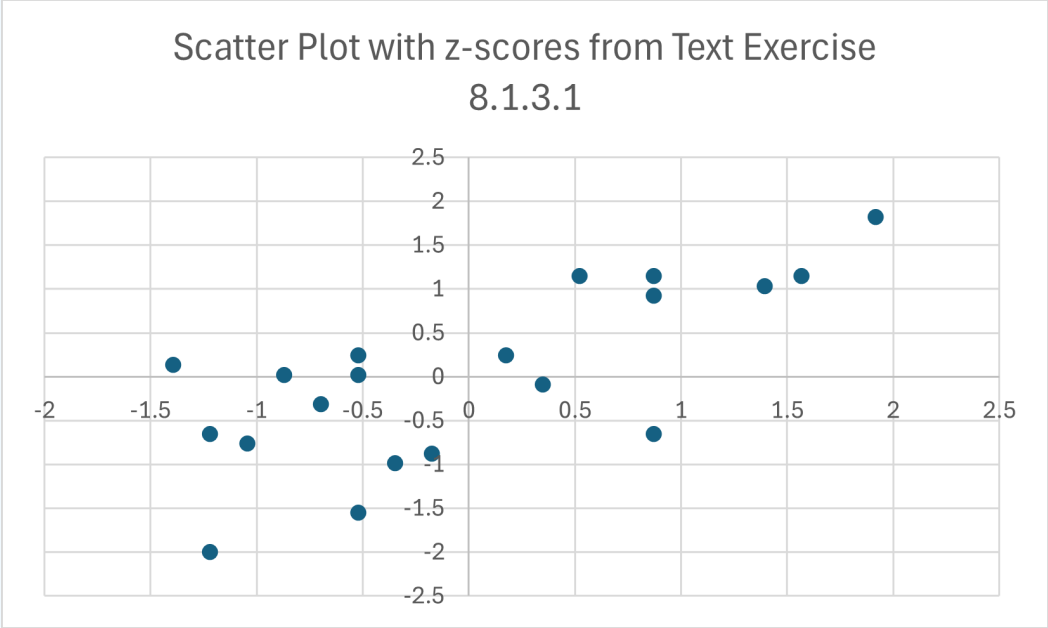


Figure 8.2.4: Scatter plot of z-scores

Answer

The scatter plot reveals points in each of the four quadrants, but the majority of the points fall in the first and third quadrants.

Note: Quadrants with Original Data

We can relate the quadrants back to the original raw data using our understanding of the z-score. Recall that a z-score is positive when the observation is larger than the mean and negative when the observation is smaller than the mean. So, the first quadrant of the scatter plot with z-scores consists of all the points where both variables were greater than the means of their respective variables. The third quadrant consists of all the points where both variables were less than the means of their respective variables, and the second and fourth quadrants consist of the points where one variable was larger and one was smaller. We can build quadrants using the means from our two variables as seen below with the data from text exercise 8.2.1.

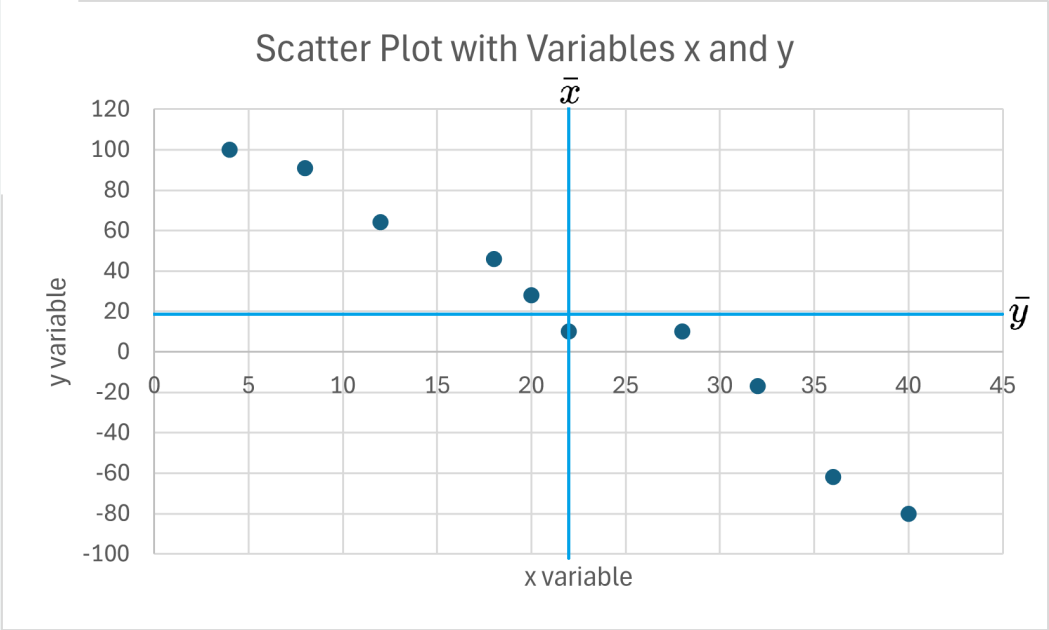


Figure 8.2.5: Scatter plot of variables x and y with $x = \bar{x}$ and $y = \bar{y}$ graphed

We begin to see that the strength of the association is related to the number of points in particular quadrants. With negative associations we see the majority of points in quadrants II and IV; while, we see the majority of points in quadrants I and III when the association is positive. Note that in quadrants I and III, the z-scores have the same sign (either both positive or both negative), and in quadrants II and IV, the z-scores have opposite signs (one negative and one positive). This means that we can readily identify how an observation contributes to the association by the sign of the product its z-scores. If the product is positive, the point falls in either quadrant I or III. If the product is negative, the point falls in either quadrant II or IV. If the majority of the products are positive, we would expect a positive association. If the majority of the products are negative, we would expect a negative association. Of course, there is more at play which we will consider now.

The points on the scatter plot of z-scores that are close to the origin could easily be seen in a path with an upward direction or a downward direction. As such, the number of points that are close to the origin in any given quadrant contribute less to the determination of the association as the points that are farther away from the origin. If the z-scores of an observation have large magnitudes, that observation contributes to the determination of the association to a larger degree. We note that both the sign and the magnitude of the product of an observation's z-scores inform us about the association between the variables. As such, it seems that a reasonable measure of linear correlation would involve averaging all of the products of each observation's z-scores. Indeed, this is what r, the Pearson correlation coefficient, does. We now provide the definition for sample, bivariate, quantitative data coming from a sample size of n with variables x and y.

$$r = \frac{1}{n-1}\sum_{i=1}^{n} z_x \cdot z_y = \frac{1}{n-1}\sum_{i=1}^{n} \frac{x_i - \bar{x}}{s_x} \cdot \frac{y_i - \bar{y}}{s_y}$$

Pearson's Correlation Coefficient

Now that we are equipped with the purpose and definition of Pearson's correlation coefficient, we will explore its meaning and implications. Perhaps the first observation made was that the multiplicative factor of $\frac{1}{n-1}$ reminds us of the adjusted averaging that takes place with sample standard deviation. This is for good reason; using such a factor guarantees that the value of r is between -1 and 1 inclusively regardless of the sample size and the magnitudes of the orignal data.

A second observation is that if we switched the variable names in the formula, the result would be exactly the same formula and compute the same value for r; the correlation between x and y is the same as the correlation between y and x. It does not matter which variable gets plotted on the horizontal axis or the vertical axis; the correlation between them will be the same. This was a desired trait mentioned earlier that is achieved with Pearson's correlation coefficient.

As previously discussed, the sign of each product in the computation of r indicates how that particular observation contributes to the association. If r is negative, the summation of the products is negative with the majority of points falling into quadrants II and IV, leading us to conclude that there is a negative correlation. If on the other hand r is positive, the summation of the products is positive which points to the majority of points falling into quadrants I and III leading us to conclude there is a positive correlation. We see that the sign of r tells us the type of correlation present between two variables thought to be linearly associated.

We are not only interested in the type of correlation, but also in the strength of the correlation. If r is 0 or close to 0, some of the products were negative while some of the products were positive; thus, when they were added together, many of the positive values cancelled out many of the negative values resulting in a very small sum. In this case, we expect the points to be distributed fairly evenly across the four quadrants, meaning, the association is weak. If $r = 0$, there is no association. If r is close to 0, there may or may not be some association. Random chance and measurement errors can lead to nonzero r values when there really is no relationship. If there is some association that produced a correlation coefficient close to 0, the weakness of the association, in general, does not warrant further interest. When the r value is 0 or close to 0, we say the variables do not exhibit any correlation.

The only time that r is equal to 1 or -1 is when the association fits a linear function perfectly: every point falls on the same line. If the slope of the line is negative, $r = -1$, and if the slope of the line is positive, $r = 1$. We can assess the strength of a linear association based on how close the value of $|r|$ is to 1.

? Text Exercise 8.2.3

1. Compute, by hand, the the Pearson's correlation coefficient r for the data set examined in text exercise 8.2.1. We replicate the table of values to facilitate the computation. Interpret the meaning of r in light of its sign and magnitude. Compare the findings with what we already know to be true.

Table 8.2.4: Table of values for variables x and y

x	y	z_x	z_y
4	100	$\frac{4-22}{12} = -\frac{3}{2}$	$\frac{100-19}{60} = \frac{27}{20}$
8	91	$\frac{8-22}{12} = -\frac{7}{6}$	$\frac{91-19}{60} = \frac{6}{5}$
12	64	$\frac{12-22}{12} = -\frac{5}{6}$	$\frac{64-19}{60} = \frac{3}{4}$
18	46	$\frac{18-22}{12} = -\frac{1}{3}$	$\frac{46-19}{60} = \frac{9}{20}$
20	28	$\frac{20-22}{12} = -\frac{1}{6}$	$\frac{18-19}{60} = \frac{3}{20}$
22	10	$\frac{22-22}{12} = 0$	$\frac{10-19}{60} = -\frac{3}{20}$
28	10	$\frac{28-22}{12} = \frac{1}{2}$	$\frac{10-19}{60} = -\frac{3}{20}$
32	-17	$\frac{32-22}{12} = \frac{5}{6}$	$\frac{-17-19}{60} = -\frac{3}{5}$
36	-62	$\frac{36-22}{12} = \frac{7}{6}$	$\frac{-62-19}{60} = -\frac{27}{20}$
40	-80	$\frac{40-22}{12} = \frac{3}{2}$	$\frac{-80-19}{60} = -\frac{33}{20}$

Answer

Table 8.2.5 Table of values and computations for variables x and y

x	y	z_x	z_y	$z_x \cdot z_y$
4	100	$\frac{4-22}{12} = -\frac{3}{2}$	$\frac{100-19}{60} = \frac{27}{20}$	$-\frac{3}{2} \cdot \frac{27}{20} = -\frac{81}{40}$
8	91	$\frac{8-22}{12} = -\frac{7}{6}$	$\frac{91-19}{60} = \frac{6}{5}$	$-\frac{7}{6} \cdot \frac{6}{5} = -\frac{7}{5}$
12	64	$\frac{12-22}{12} = -\frac{5}{6}$	$\frac{64-19}{60} = \frac{3}{4}$	$-\frac{5}{6} \cdot \frac{3}{4} = -\frac{5}{8}$
18	46	$\frac{18-22}{12} = -\frac{1}{3}$	$\frac{46-19}{60} = \frac{9}{20}$	$-\frac{1}{3} \cdot \frac{9}{20} = -\frac{3}{20}$
20	28	$\frac{20-22}{12} = -\frac{1}{6}$	$\frac{18-19}{60} = \frac{3}{20}$	$-\frac{1}{6} \cdot \frac{3}{20} = -\frac{1}{40}$
22	10	$\frac{22-22}{12} = 0$	$\frac{10-19}{60} = -\frac{3}{20}$	$0 \cdot -\frac{3}{20} = 0$
28	10	$\frac{28-22}{12} = \frac{1}{2}$	$\frac{10-19}{60} = -\frac{3}{20}$	$\frac{1}{2} \cdot -\frac{3}{20} = -\frac{3}{40}$
32	-17	$\frac{32-22}{12} = \frac{5}{6}$	$\frac{-17-19}{60} = -\frac{3}{5}$	$\frac{5}{6} \cdot -\frac{3}{5} = -\frac{1}{2}$
36	-62	$\frac{36-22}{12} = \frac{7}{6}$	$\frac{-62-19}{60} = -\frac{27}{20}$	$\frac{7}{6} \cdot -\frac{27}{20} = -\frac{63}{40}$
40	-80	$\frac{40-22}{12} = \frac{3}{2}$	$\frac{-80-19}{60} = -\frac{33}{20}$	$\frac{3}{2} \cdot -\frac{33}{20} = -\frac{99}{40}$

Having computed all of the products of the z-scores in the far right column, all we need to do is sum them up and divide by 9.

$$r = \frac{1}{10-1} \sum_{i=1}^{10} \frac{x_i - \bar{x}}{s_x} \cdot \frac{y_i - \bar{y}}{s_y} = \frac{1}{9} \cdot -\frac{177}{20} = -\frac{59}{60} = -0.98\bar{3}$$

The negative value of r indicates a negative correlation between the variables x and y. The magnitude of 0.9833which is close to 1 indicates that the linear association between x and y is quite strong. The understanding derived from our considerations of the correlation coefficient r align with the conclusions previously drawn.

2. Consider the following bivariate data set using both a scatter plot and r.

Table 8.2.6: Table of values for variables x and y

x	y
-2	4
0	4
1	4
2	4
5	4

Answer

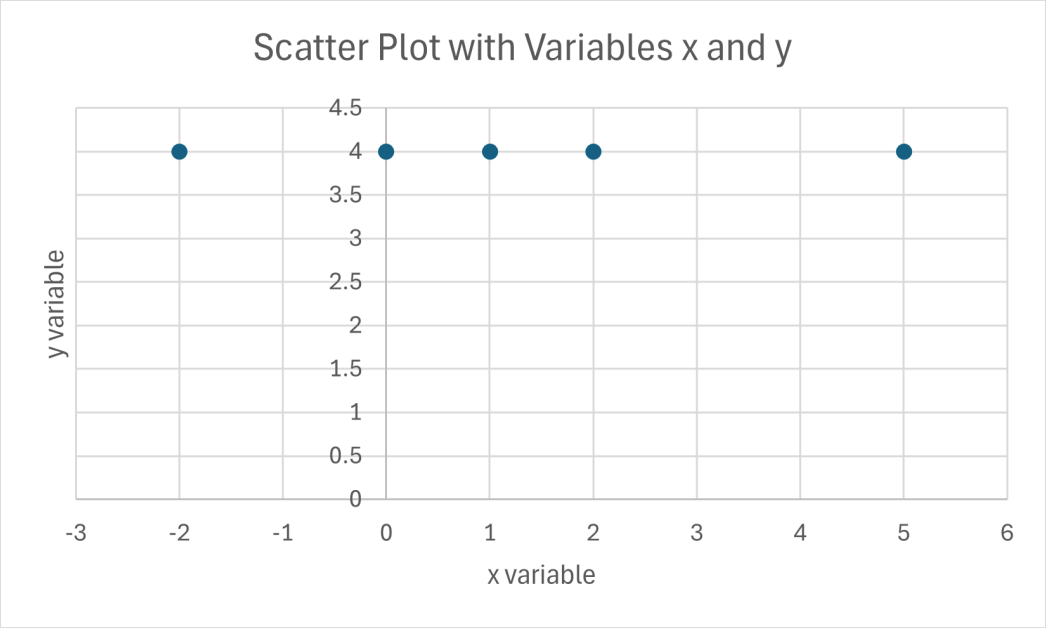


Figure 8.2.6 Scatter plot

The scatter plot shows points that form a perfectly straight horizontal line. When computing the correlation coefficient, we run into problems, because the standard deviation of the variable y is 0. We cannot divide by 0. Since every y value is 4, the mean is 4, and the standard deviation is 0. As such, it appears that the correlation coefficient fails to recognize a perfectly linear association. In reality this is not the case. When our data aligns in either a perfect horizontal line or vertical line, there is no association between the variables. There is no relationship between the value of the x variable and the value of the y variable because one of the variables is fixed regardless of the other variable. Since there is no relationship between the variables, it does not make sense to measure the degree of a linear relationship.

As we saw with a data set containing just 10 observations, the computation of the correlation coefficient can be quite tedious. As such, the computation is often carried out using technology. The function within Excel that computes Pearson's correlation coefficient r is the CORREL function. The function takes, as inputs, two arrays of numbers that are the same size. The first array consists of all the values of one of the variables; the second array consists of all the values of the other variable. The program matches the arrays by position to pair the values of each variable.

3. Using technology, reconstruct the scatter plot of the previously considered data and calculate Pearson's correlation coefficient r.

Table 8.2.7: Age of bride and room on wedding day

Groom's Age (years)	Bride's Age (years)
20	21
26	20
32	34
30	30
21	22
29	28
26	25
34	34
29	28

Groom's Age (years)	Bride's Age (years)
55	50
30	26
43	39
30	29
24	22
20	19

Answer

$r = 0.9689$

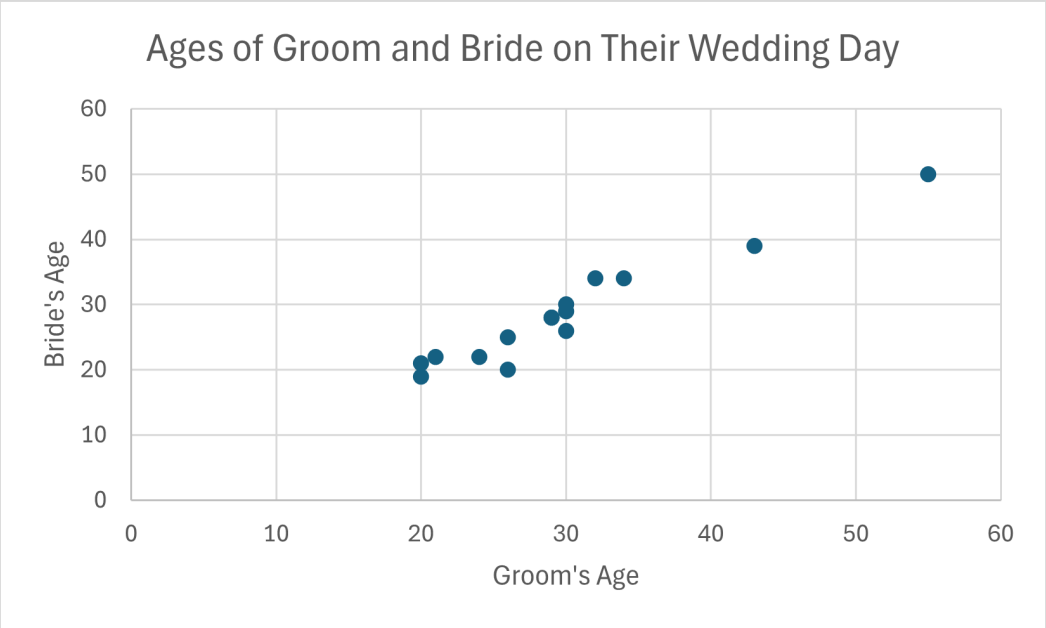


Figure 8.2.7: Scatter plot of ages of bride and groom

4. Using technology, compute Pearson's correlation coefficient r.

Table 8.2.8: Age of bride vs number of children

Bride's Age (years)	Number of Children
19	2
21	8
21	5
22	0
22	3
23	6
23	4
23	2
24	7

Bride's Age (years)	Number of Children
25	1
25	3
26	3
27	4
29	5
31	3
31	1
32	2
33	0
35	2
41	1

Answer

$r = -0.4247$

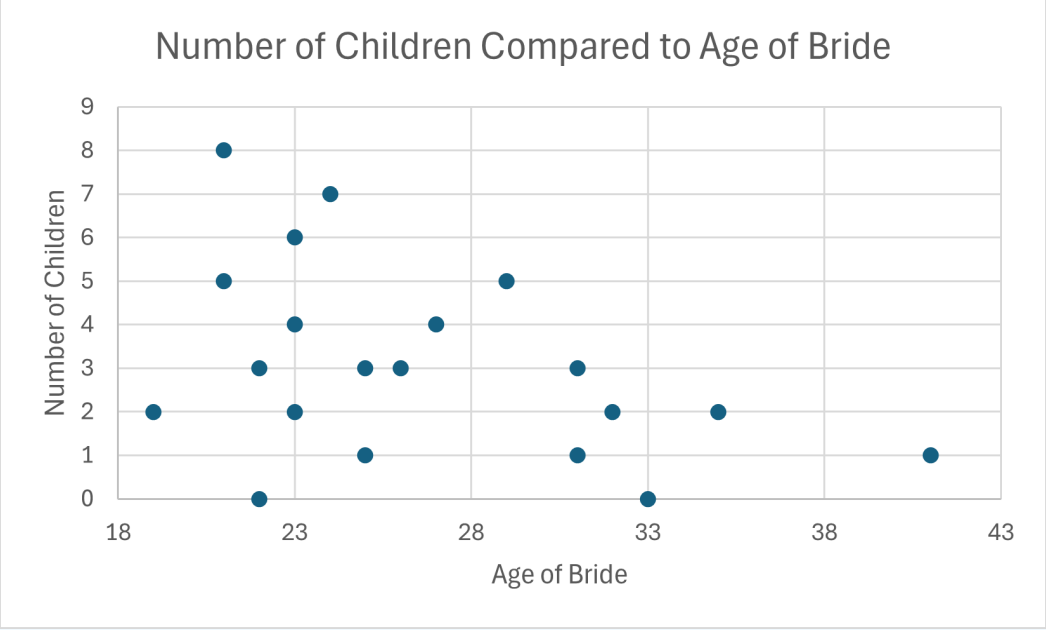


Figure 8.2.8 Scatter plot of age of bride and number of children

5. Using technology, compute Pearson's correlation coefficient r.

Table 8.2.9: Bivariate data from Text Exercise 8.1.3.1

x	y
4	5
11	4.5
10	6
14	10

x	y
6	6
6	-2
2	2
8	1
12	10
7	0.5
14	9
5	3.5
3	1.5
17	9.5
1	5.5
18	10
14	2
20	13
2	-4
6	5

Answer

$r = 0.7253$

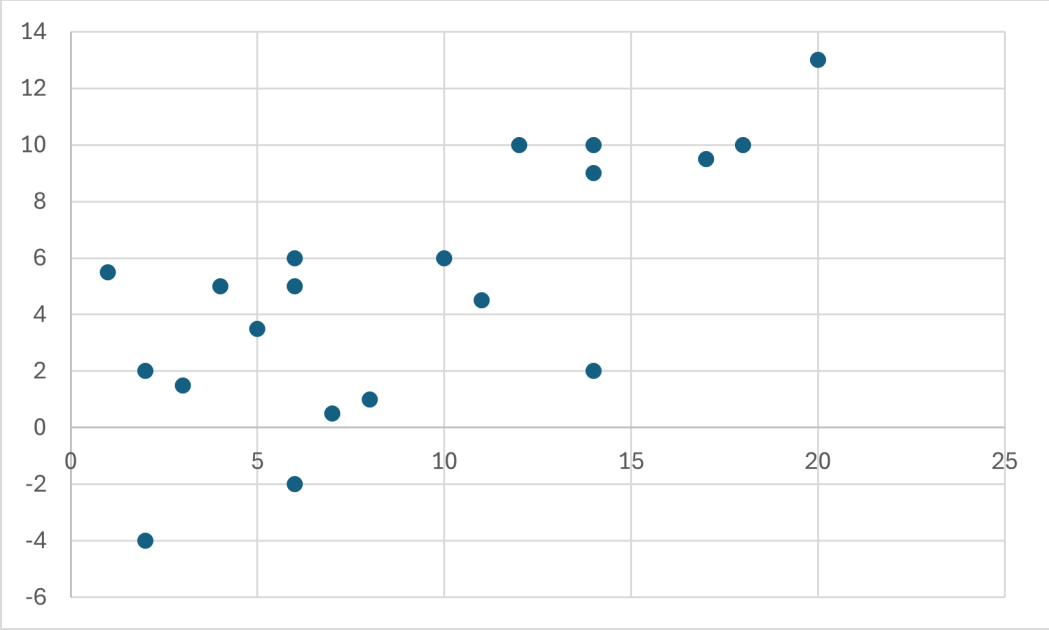


Figure 8.2.9 Scatter plot

6. Using technology, compute Pearson's correlation coefficient r.

Table 8.2.10: Bivariate data from Text Exercise 8.1.3.4

x	y
12	-6
10	-2
12	-6
6	9
20	-19
15	-8
17	-15
3	14
6	8
18	-15
5	10
1	20
12	-4
6	8
2	14
20	-19
6	8
4	10
4	12
13	-5

Answer

$r = -0.9941$

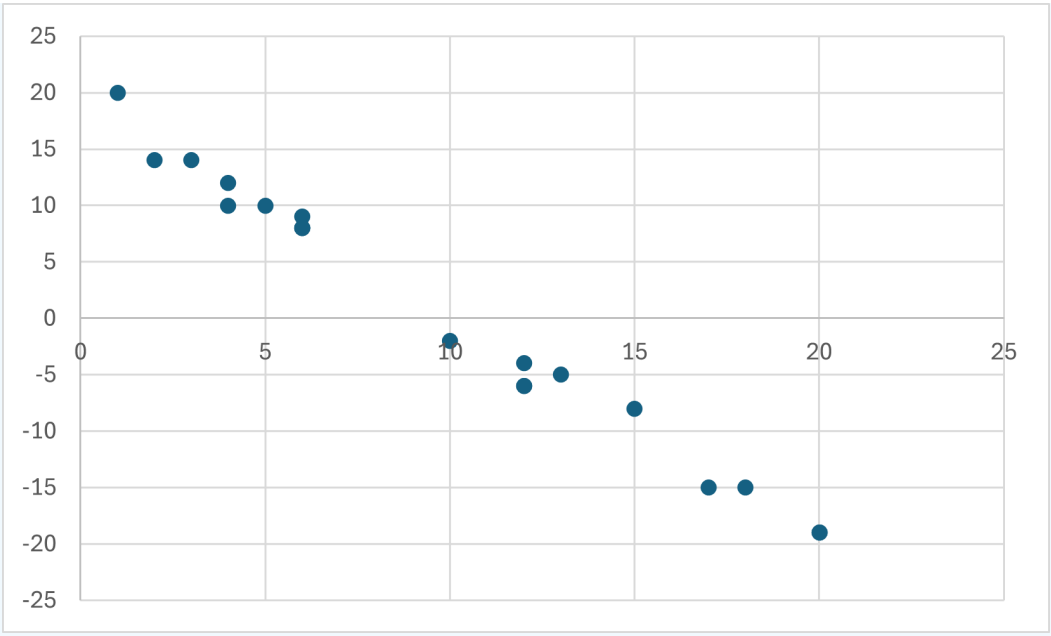


Figure 8.2.10 Scatter plot

Text Exercise 8.2.4

For each of the following scatter plots, determine if the proposed r value is reasonable. If not, explain why.

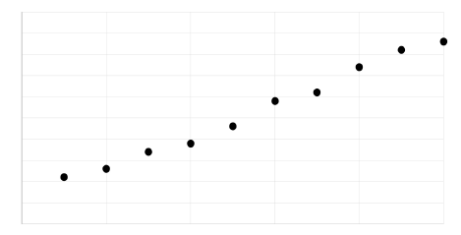
1. 

Figure 8.2.11: Scatter plot

$r = 0.99$

Answer

The scatter plot displays strong positive correlation that appears very close to a line. As such, a positive value close to 1 seems reasonable.

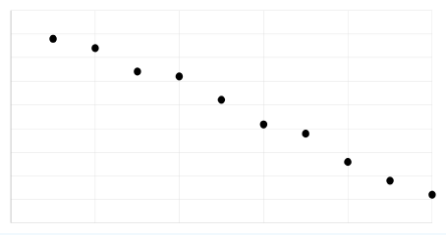
2. 

Figure 8.2.12: Scatter plot

$r = 0.95$

Answer

The scatter plot displays strong negative correlation that appears very close to a line. As such, the proposed r value is unreasonable. The magnitude might be reasonable, but it is clear that the r value should be negative.

3. 

Figure 8.2.13: Scatter plot

$$r = 0.85$$

Answer

The scatter plot does not display much correlation. As such, a value of 0.85 seems unreasonably high.

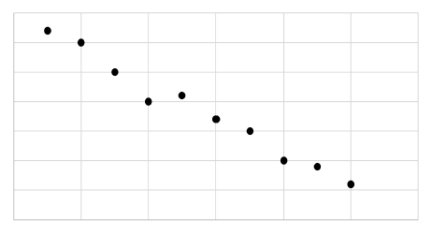
4. 

Figure 8.2.14: Scatter plot

$$r = -1$$

Answer

The scatter plot displays a strong negative linear correlation. However, we can easily see that the data does not fit perfectly on a line. As such, the proposed r value is unreasonable.

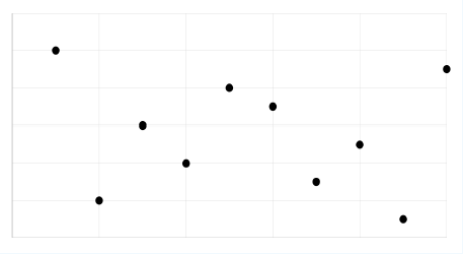
5. 

Figure 8.2.15: Scatter plot

$$r = -0.15$$

Answer

The scatter plot does not display much correlation. As such, an r value close to 0 seems reasonable. Given the general downward direction of the data, a negative r value seems reasonable. We thus conclude that such an r value seems reasonable.

Text Exercise 8.2.5

1. The statistician Francis Anscombe constructed 4 data sets in 1973 that have earned the moniker Anscombe's Quartet. In this text exercise, we will analyze each of the data sets individually and then consider them together. For each of the data sets, compute $\bar{x}$, s_x, $\bar{y}$, s_y, and r. What similarities are there between the data sets? What conclusions can be drawn?

Table 8.2.11: Anscombe's Quartet

Data Set I (x,y)		Data Set II (x,y)		Data Set III (x,y)		Data Set IV (x,y)	
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Answer

Data Set I: $\bar{x} = 9$, $s_x \approx 3.3166$,$\bar{y} \approx 7.5009$,$s_y \approx 2.0316$,and $r \approx 0.8164$

Data Set II: $\bar{x} = 9$, $s_x \approx 3.3166$,$\bar{y} \approx 7.5009$,$s_y \approx 2.0317$,and $r \approx 0.8162$

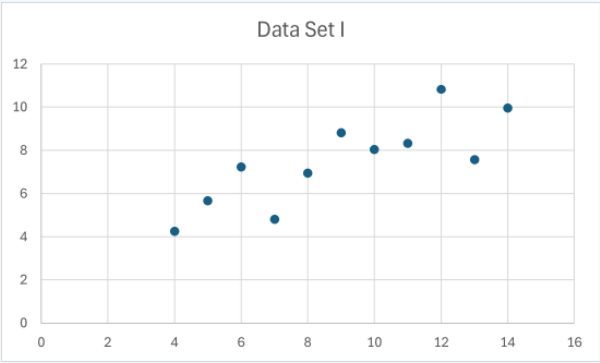
Data Set III: $\bar{x} = 9$, $s_x \approx 3.3166$,$\bar{y} \approx 7.5$, $s_y \approx 2.0304$,and $r \approx 0.8163$

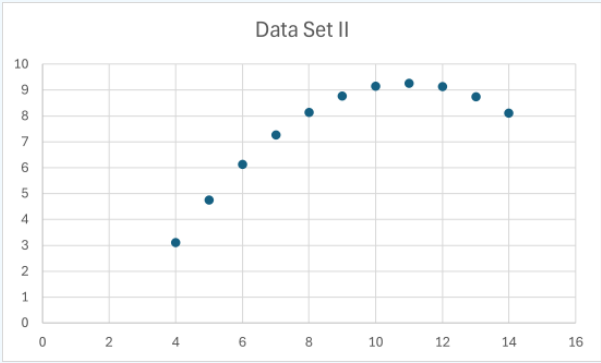
Data Set IV: $\bar{x} = 9$, $s_x \approx 3.3166$,$\bar{y} \approx 7.5009$,$s_y \approx 2.0306$,and $r \approx 0.8165$

The summary statistics for each of the data sets are remarkably similar. They all match up to two or three decimal places. From the perspective of these summary statistics, the data sets are almost indistinguishable.

2. Having computed the summary statistics for each data set, construct scatter plots for each of the data sets. What conclusions can now be drawn? What implications does this exercise have on conducting statistical analyses?

Answer





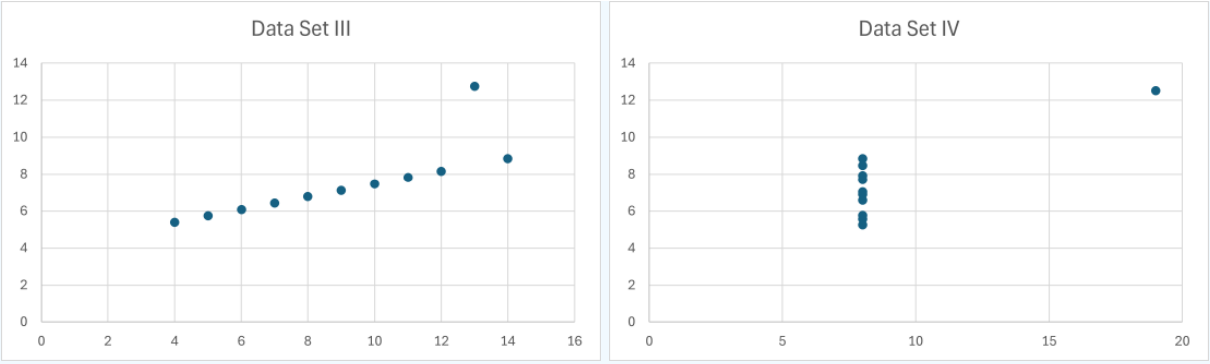


Figure 8.2.16 Scatter plots of Anscombe's Quartet

The scatter plots provide crucial information regarding the various data sets. The first scatter plot reveals an apparent linear relationship with a fairly decent association (the r value confirms this). The second scatter plot reveals a nonlinear relationship indicating that using the correlation coefficient to describe the relationship should not have been done. Indeed, for this particular nonlinear relationship, it appears for some values the association is positive, but for others the association is negative. Thus measuring for any correlation, linear or not, is not to be done. The third scatter plot reveals a quite strong linear relationship with an apparent outlier which resulted in the correlation coefficient dropping significantly in value. The fourth scatter plot reveals the presence of an outlier. The remainder of the data seems to indicate that there is no relationship between the variables.

The visual representations of data are crucial components of statistical analysis. The scatter plots provide us with the information to determine if it is even reasonable to use the correlation coefficient as a measure. When conducting statistical analyses, constructing visualizations is generally the first step, as they provide much information and a general feel for the data. A final takeaway is that we do not want to blindly compute the correlation coefficient and use it as a measure for the presence of a linear relationship. It is best used as a way to measure the strength of a linear relationship that is thought to exist based on scatter plots or other lines of reasoning.
