Chapter 10

Correlation and Regression

Chapter 10 Overview

Introduction

- 10-1 Scatter Plots and Correlation
- 10-2 Regression
- 10-3 Coefficient of Determination and Standard Error of the Estimate
- 10-4 Multiple Regression (Optional)

Chapter 10 Objectives

1. Draw a scatter plot for a set of ordered pairs.
2. Compute the correlation coefficient.
3. Test the hypothesis H_0: $\rho = 0$.
4. Compute the equation of the regression line.
5. Compute the coefficient of determination.
6. Compute the standard error of the estimate.
7. Find a prediction interval.
8. Be familiar with the concept of multiple regression.

Introduction

Introduction

- In addition to hypothesis testing and confidence intervals, inferential statistics involves determining whether a relationship between two or more numerical or quantitative variables exists.

Introduction

Introduction

- **Correlation** is a statistical method used to determine whether a linear relationship between variables exists.

- **Regression** is a statistical method used to describe the nature of the relationship between variables—that is, positive or negative, linear or nonlinear.

Introduction

Introduction

- The purpose of this chapter is to answer these questions statistically:

  1. Are two or more variables related?
  2. If so, what is the strength of the relationship?
  3. What type of relationship exists?
  4. What kind of predictions can be made from the relationship?

Introduction

Introduction

1. Are two or more variables related?

2. If so, what is the strength of the relationship?

To answer these two questions, statisticians use the **correlation coefficient**, a numerical measure to determine whether two or more variables are related and to determine the strength of the relationship between or among the variables.

Introduction

3. What type of relationship exists?

Introduction

3. What type of relationship exists?

There are two types of relationships: simple and multiple.

Introduction

3. What type of relationship exists?

There are two types of relationships: simple and multiple.

Introduction

3. What type of relationship exists?

There are two types of relationships: simple and multiple.

In a simple relationship, there are two variables: an **independent variable** (predictor variable) and a **dependent variable** (response variable).

Introduction

3. What type of relationship exists?

There are two types of relationships: simple and multiple.

In a simple relationship, there are two variables: an **independent variable** (predictor variable) and a **dependent variable** (response variable).

Introduction

3. What type of relationship exists?

There are two types of relationships: simple and multiple.

In a simple relationship, there are two variables: an **independent variable** (predictor variable) and a **dependent variable** (response variable).

In a multiple relationship, there are two or more independent variables that are used to predict one dependent variable.

Introduction

Introduction

4. What kind of predictions can be made from the relationship?

Predictions are made in all areas and daily. Examples include weather forecasting, stock market analyses, sales predictions, crop predictions, gasoline price predictions, and sports predictions. Some predictions are more accurate than others, due to the strength of the relationship. That is, the stronger the relationship is between variables, the more accurate the prediction is.

10.1 Scatter Plots and Correlation

10.1 Scatter Plots and Correlation

- A **scatter plot** is a graph of the ordered pairs (*x*, *y*) of numbers consisting of the independent variable *x* and the dependent variable *y*.

Chapter 10
Correlation and Regression

Section 10-1

Example 10-1

Page #536

Example 10-1: Car Rental Companies

Construct a scatter plot for the data shown for car rental companies in the United States for a recent year.

Company	Cars (in ten thousands)	Revenue (in billions)
A	63.0	$7.0
B	29.0	3.9
C	20.8	2.1
D	19.1	2.8
E	13.4	1.4
F	8.5	1.5

Example 10-1: Car Rental Companies

Construct a scatter plot for the data shown for car rental companies in the United States for a recent year.

Company	Cars (in ten thousands)	Revenue (in billions)
A	63.0	$7.0
B	29.0	3.9
C	20.8	2.1
D	19.1	2.8
E	13.4	1.4
F	8.5	1.5

Step 1: Draw and label the *x* and *y* axes.

Example 10-1: Car Rental Companies

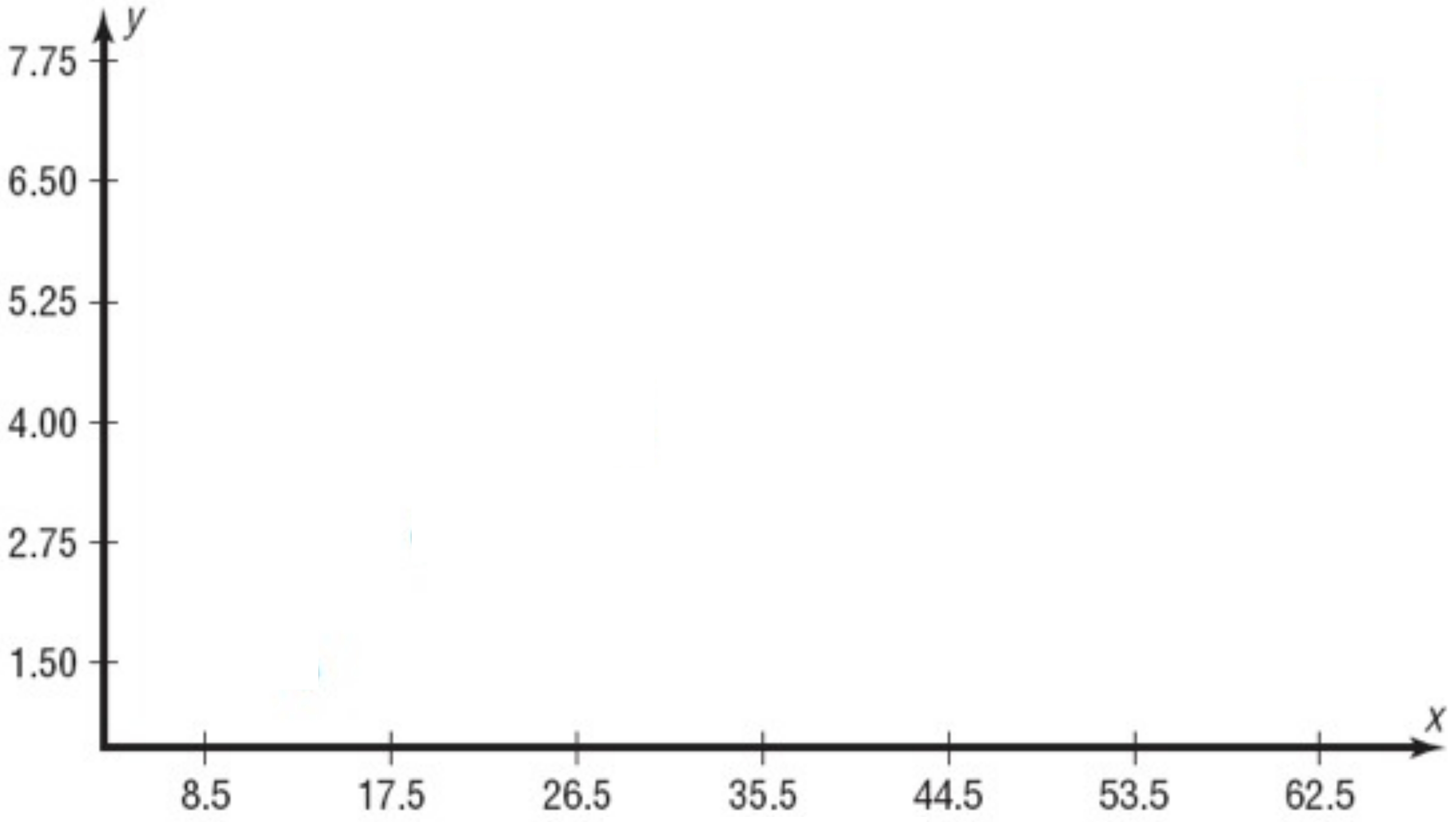
Construct a scatter plot for the data shown for car rental companies in the United States for a recent year.

Company	Cars (in ten thousands)	Revenue (in billions)
A	63.0	$7.0
B	29.0	3.9
C	20.8	2.1
D	19.1	2.8
E	13.4	1.4
F	8.5	1.5

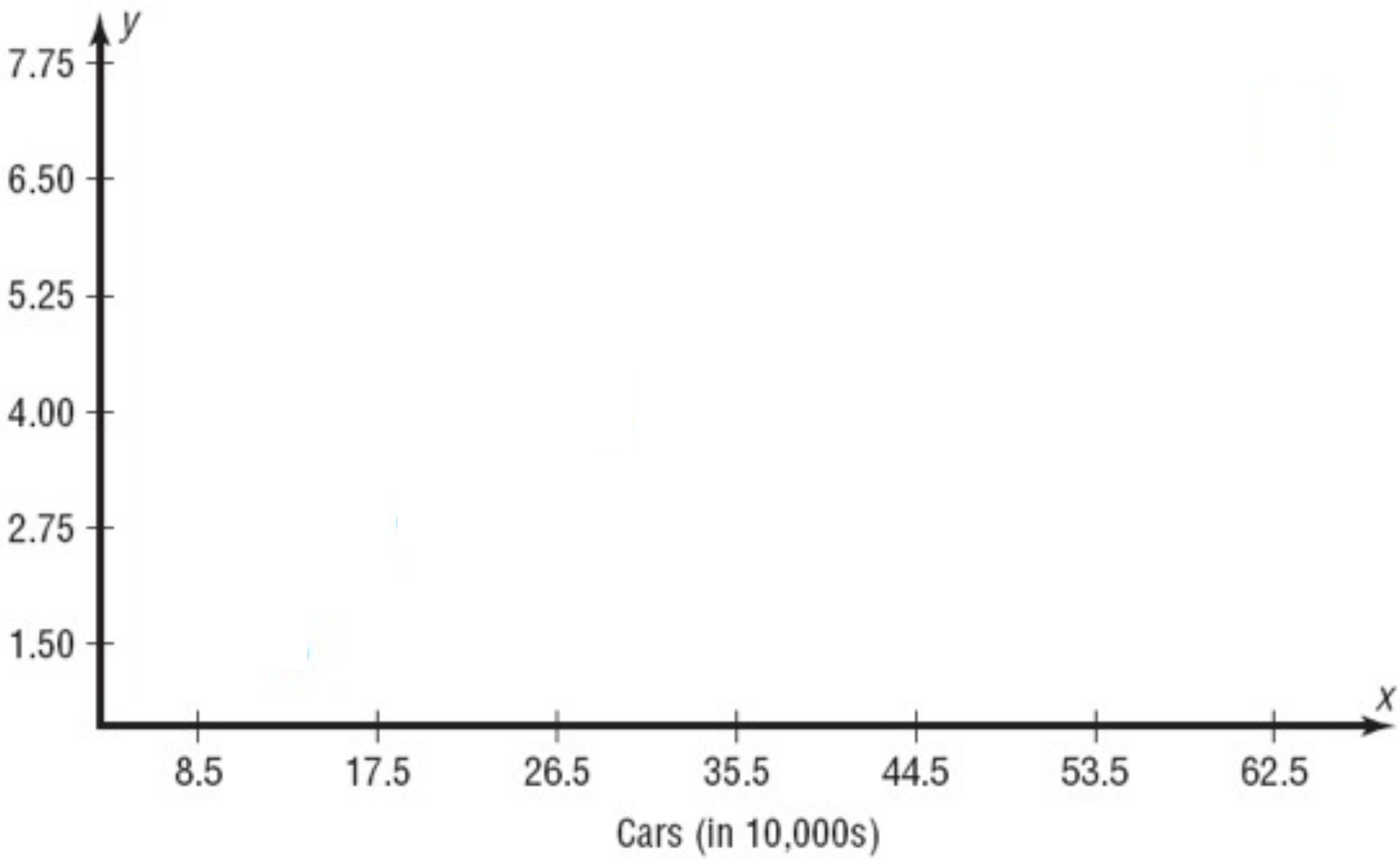
Step 1: Draw and label the *x* and *y* axes.

Step 2: Plot each point on the graph.

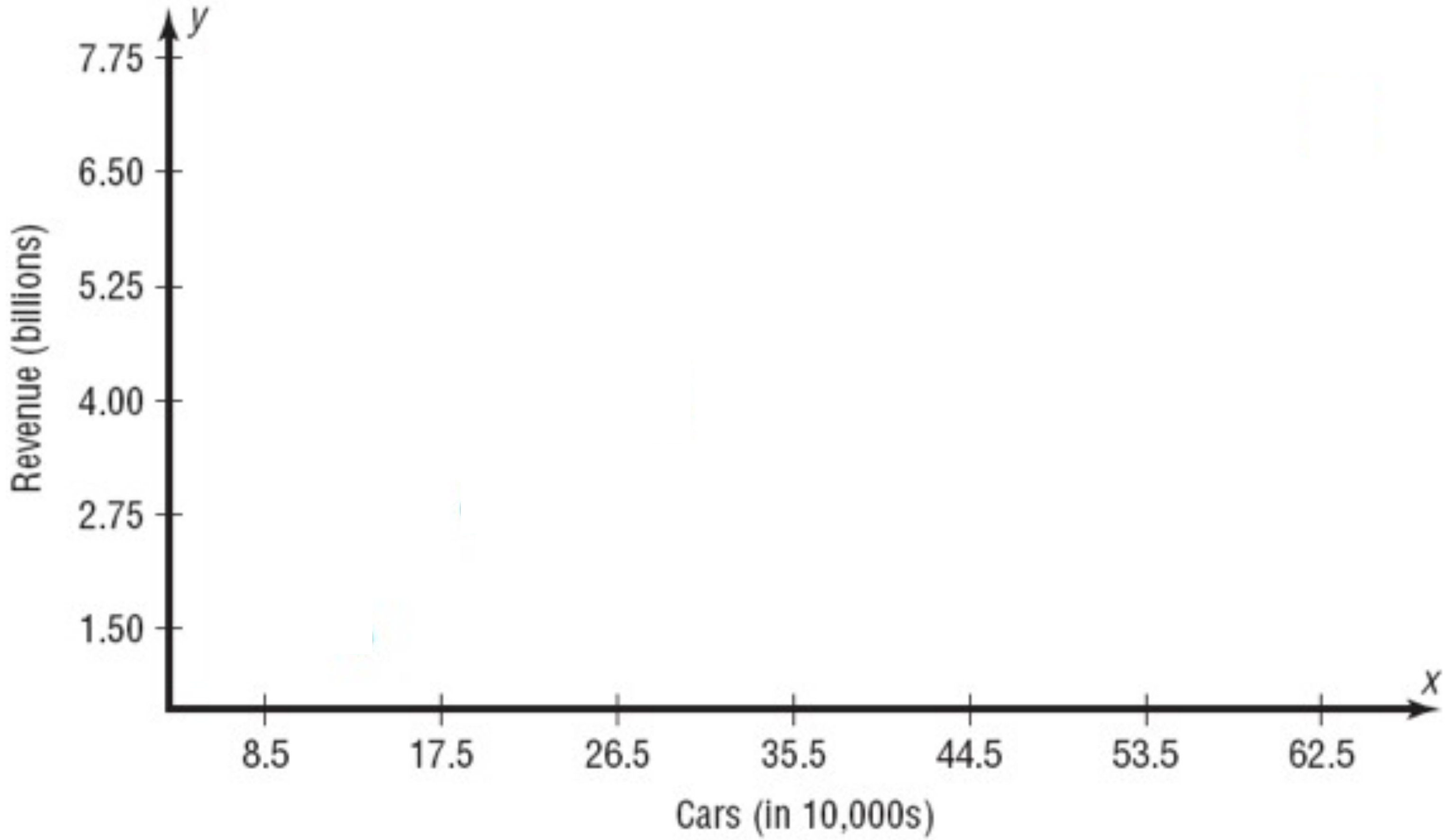
Example 10-1: Car Rental Companies



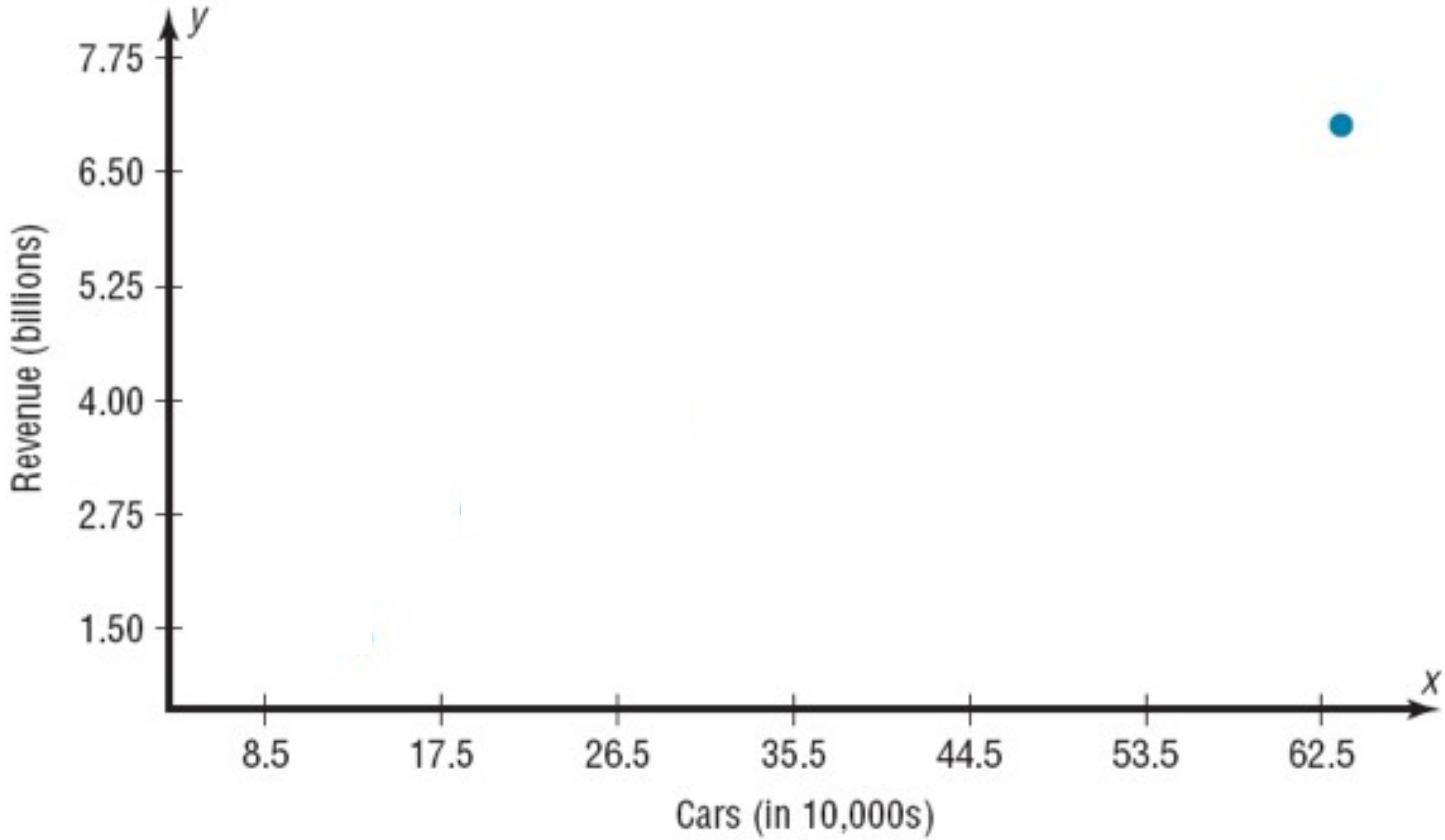
Example 10-1: Car Rental Companies



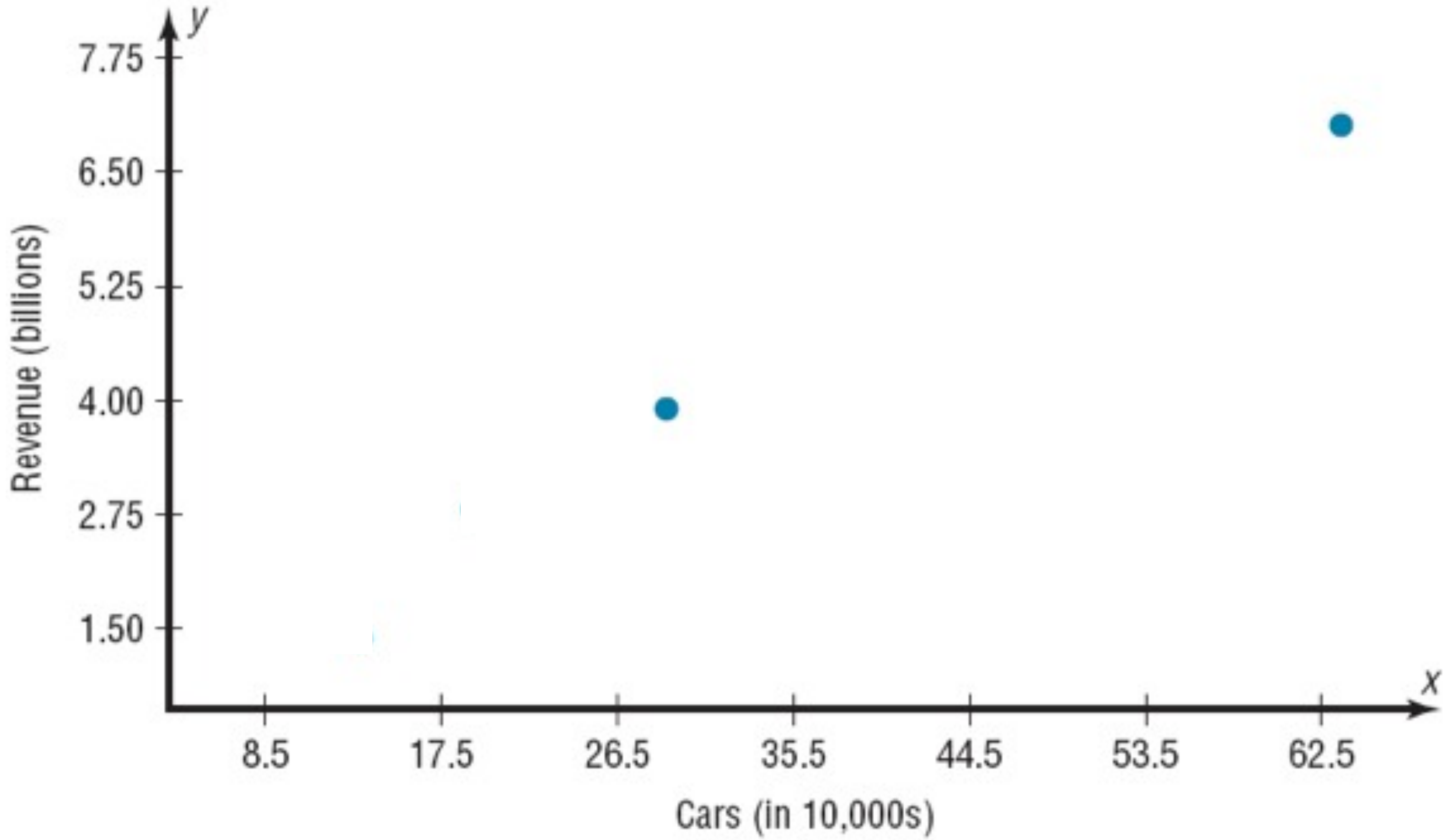
Example 10-1: Car Rental Companies



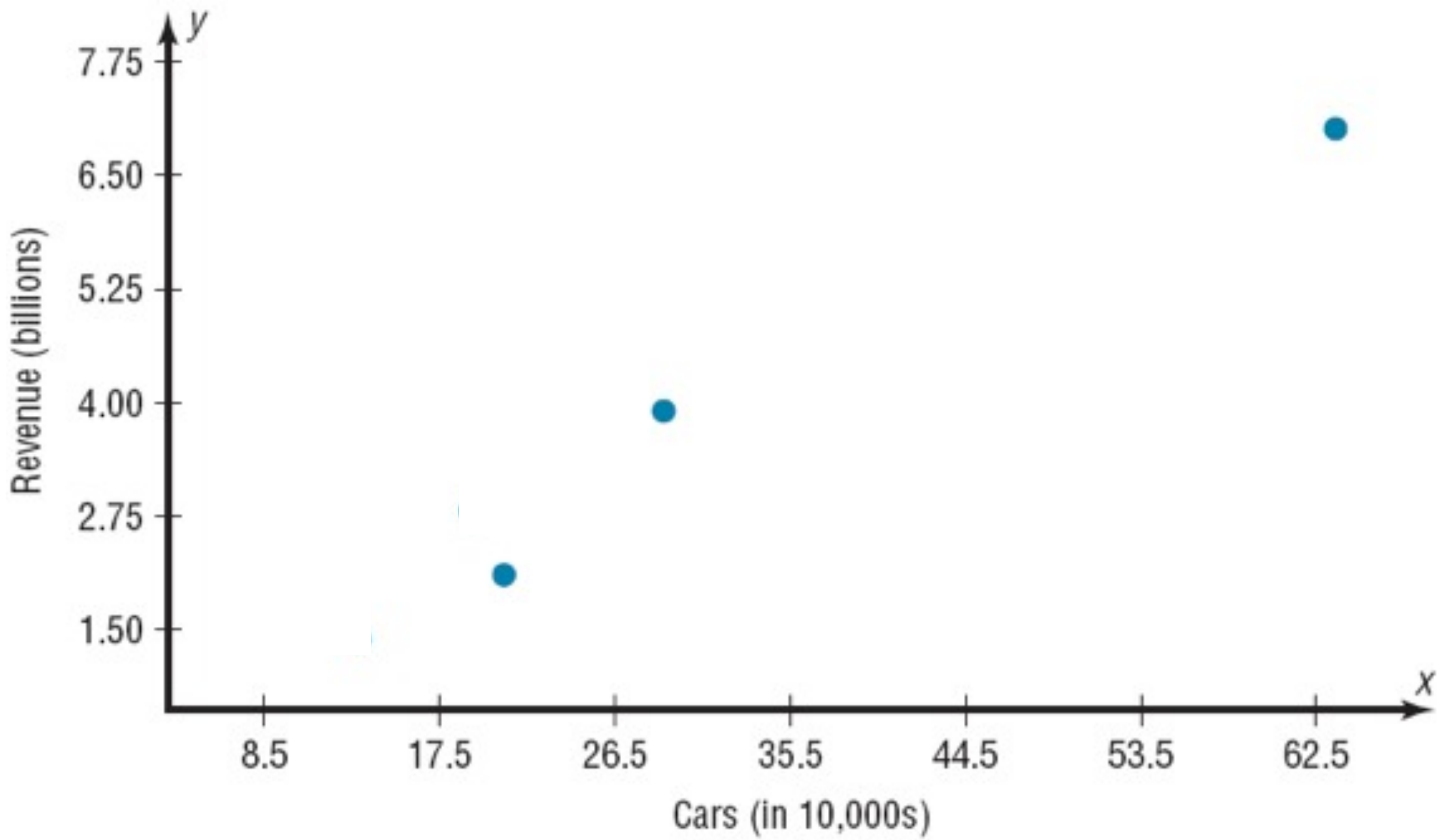
Example 10-1: Car Rental Companies



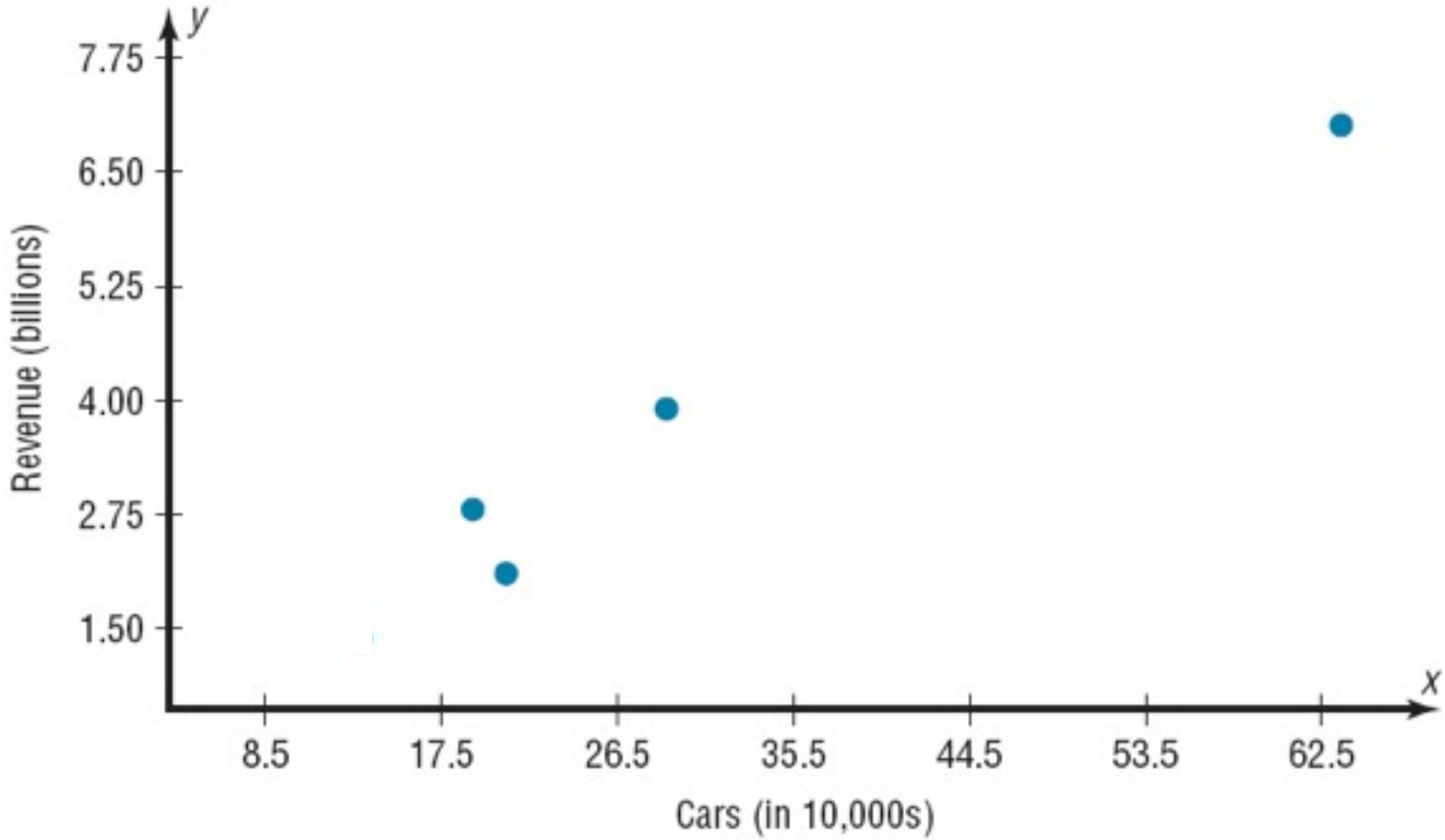
Example 10-1: Car Rental Companies



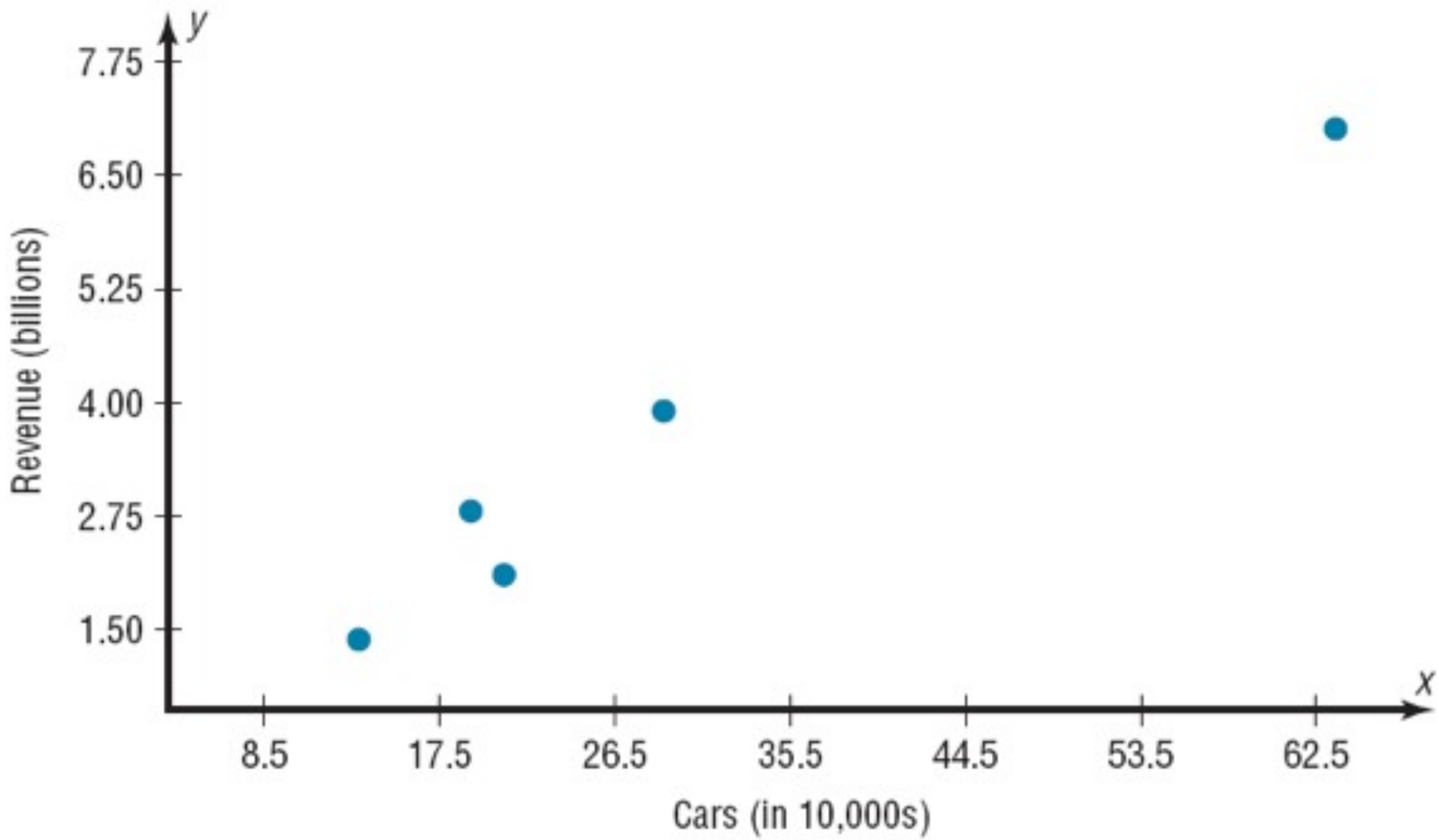
Example 10-1: Car Rental Companies



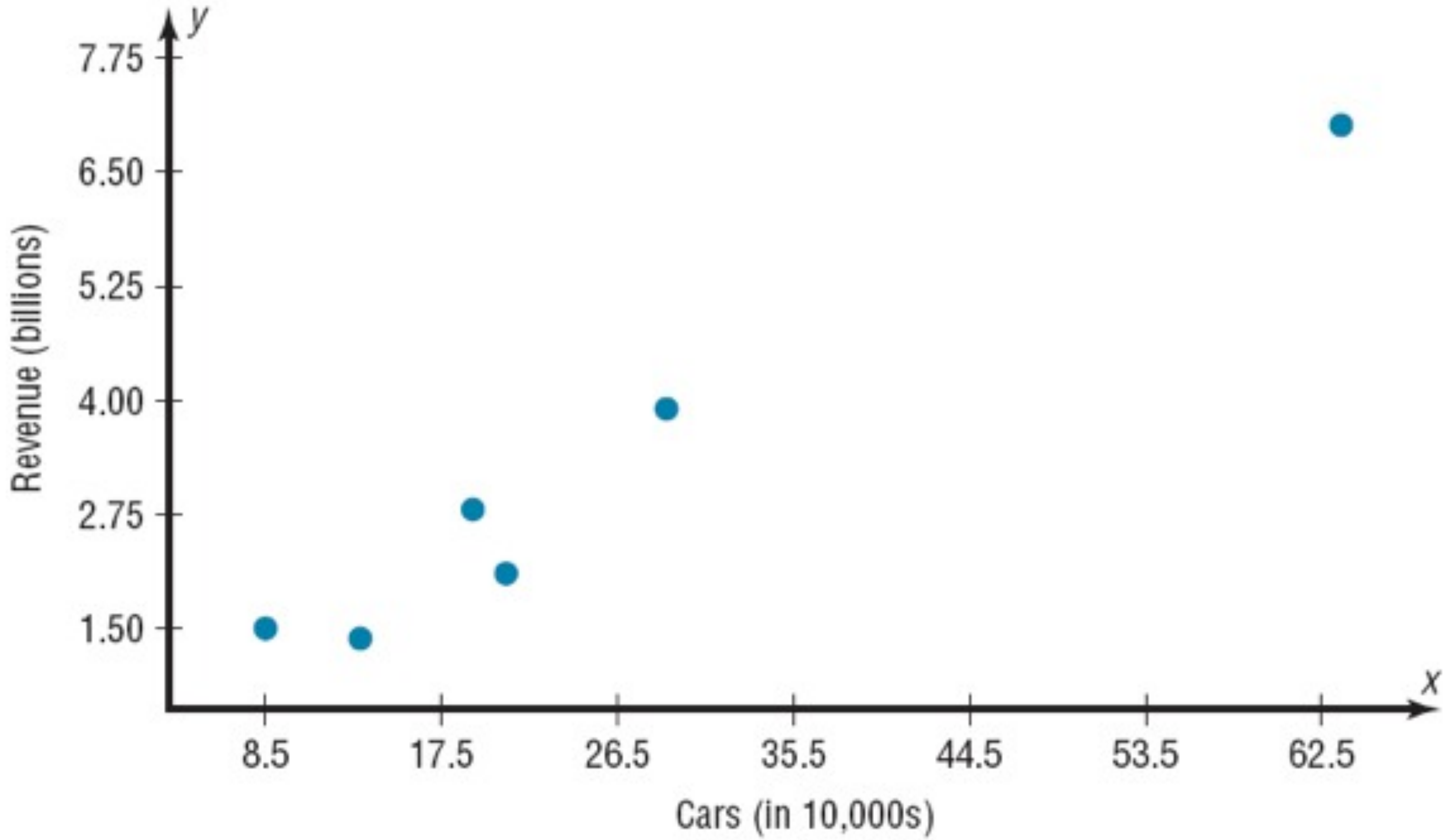
Example 10-1: Car Rental Companies



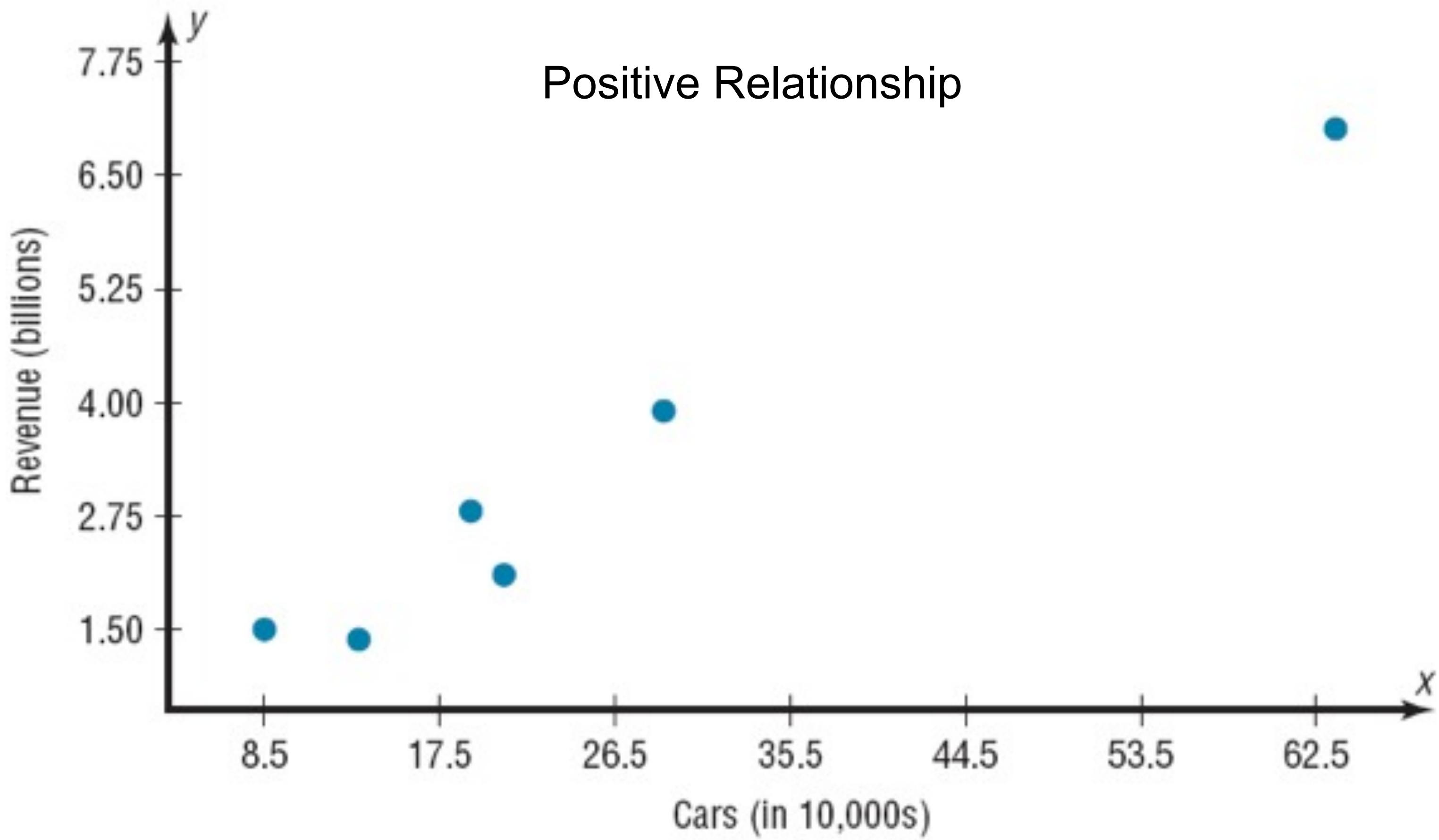
Example 10-1: Car Rental Companies



Example 10-1: Car Rental Companies



Example 10-1: Car Rental Companies



Chapter 10
Correlation and Regression

Section 10-1

Example 10-2

Page #537

Example 10-2: Absences/Final Grades

Construct a scatter plot for the data obtained in a study on the number of absences and the final grades of seven randomly selected students from a statistics class.

Student	Number of absences x	Final grade y (%)
A	6	82
B	2	86
C	15	43
D	9	74
E	12	58
F	5	90
G	8	78

Example 10-2: Absences/Final Grades

Construct a scatter plot for the data obtained in a study on the number of absences and the final grades of seven randomly selected students from a statistics class.

Student	Number of absences x	Final grade y (%)
A	6	82
B	2	86
C	15	43
D	9	74
E	12	58
F	5	90
G	8	78

Step 1: Draw and label the x and y axes.

Example 10-2: Absences/Final Grades

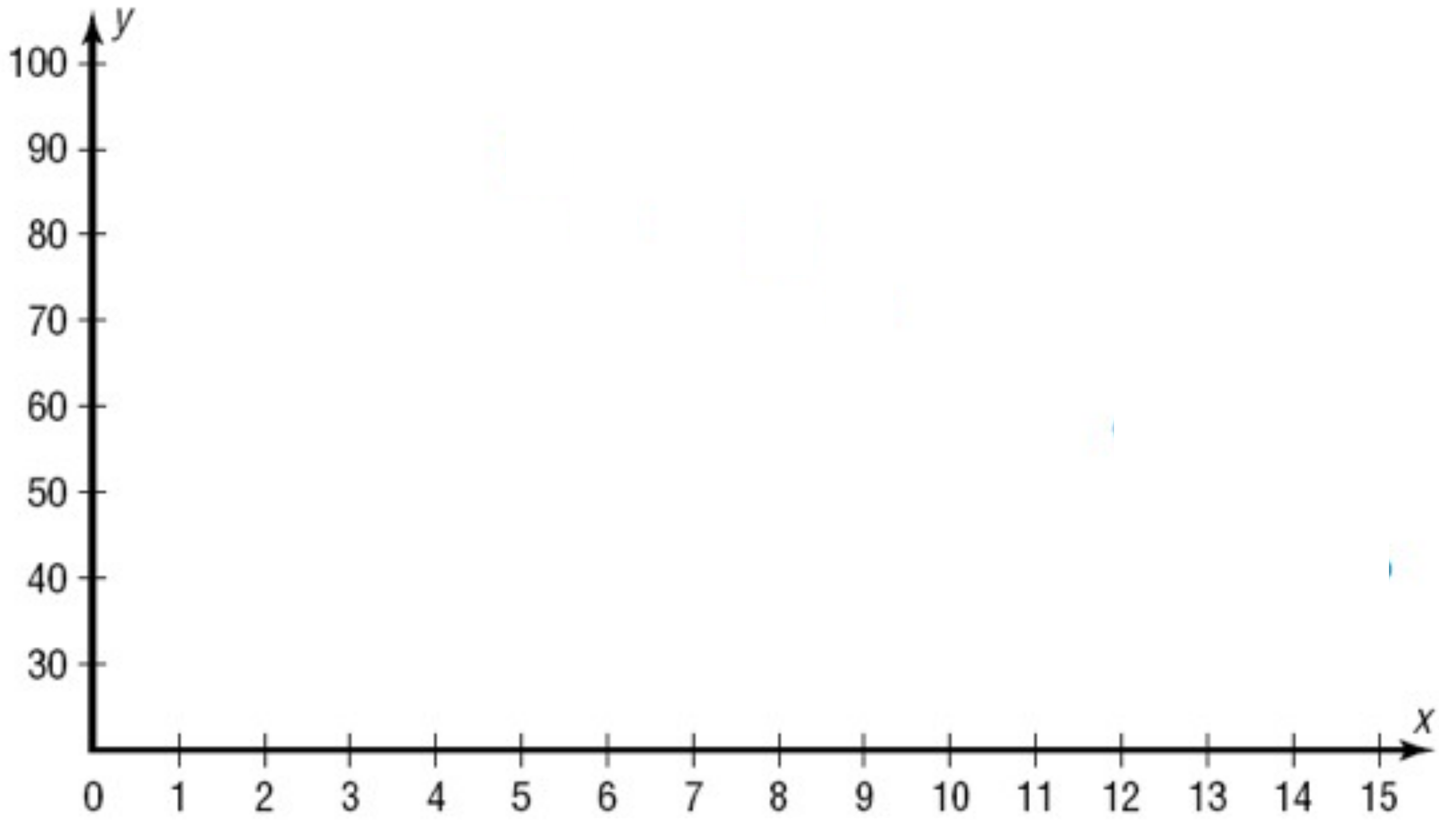
Construct a scatter plot for the data obtained in a study on the number of absences and the final grades of seven randomly selected students from a statistics class.

Student	Number of absences *x*	Final grade *y* (%)
A	6	82
B	2	86
C	15	43
D	9	74
E	12	58
F	5	90
G	8	78

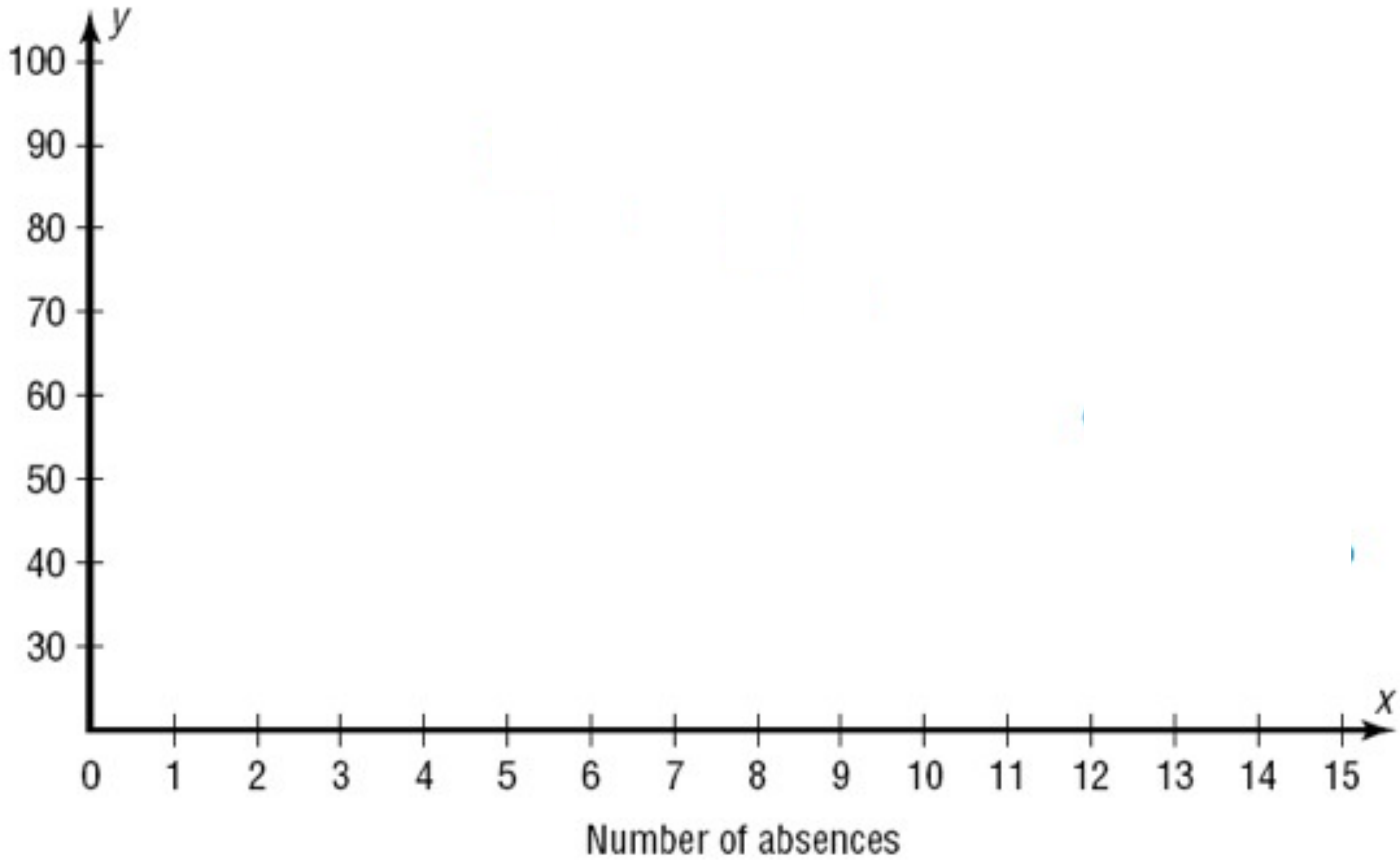
Step 1: Draw and label the *x* and *y* axes.

Step 2: Plot each point on the graph.

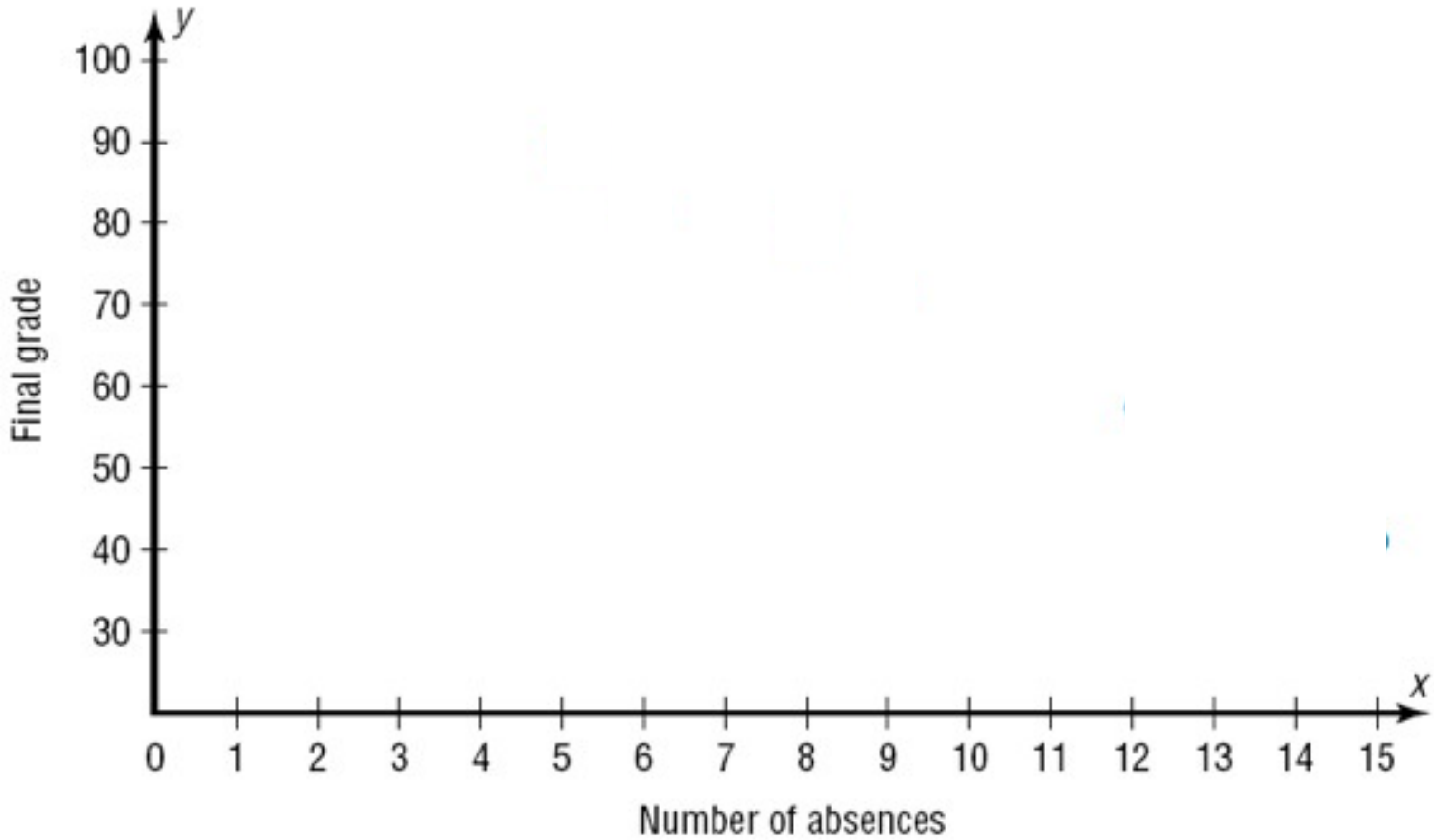
Example 10-2: Absences/Final Grades



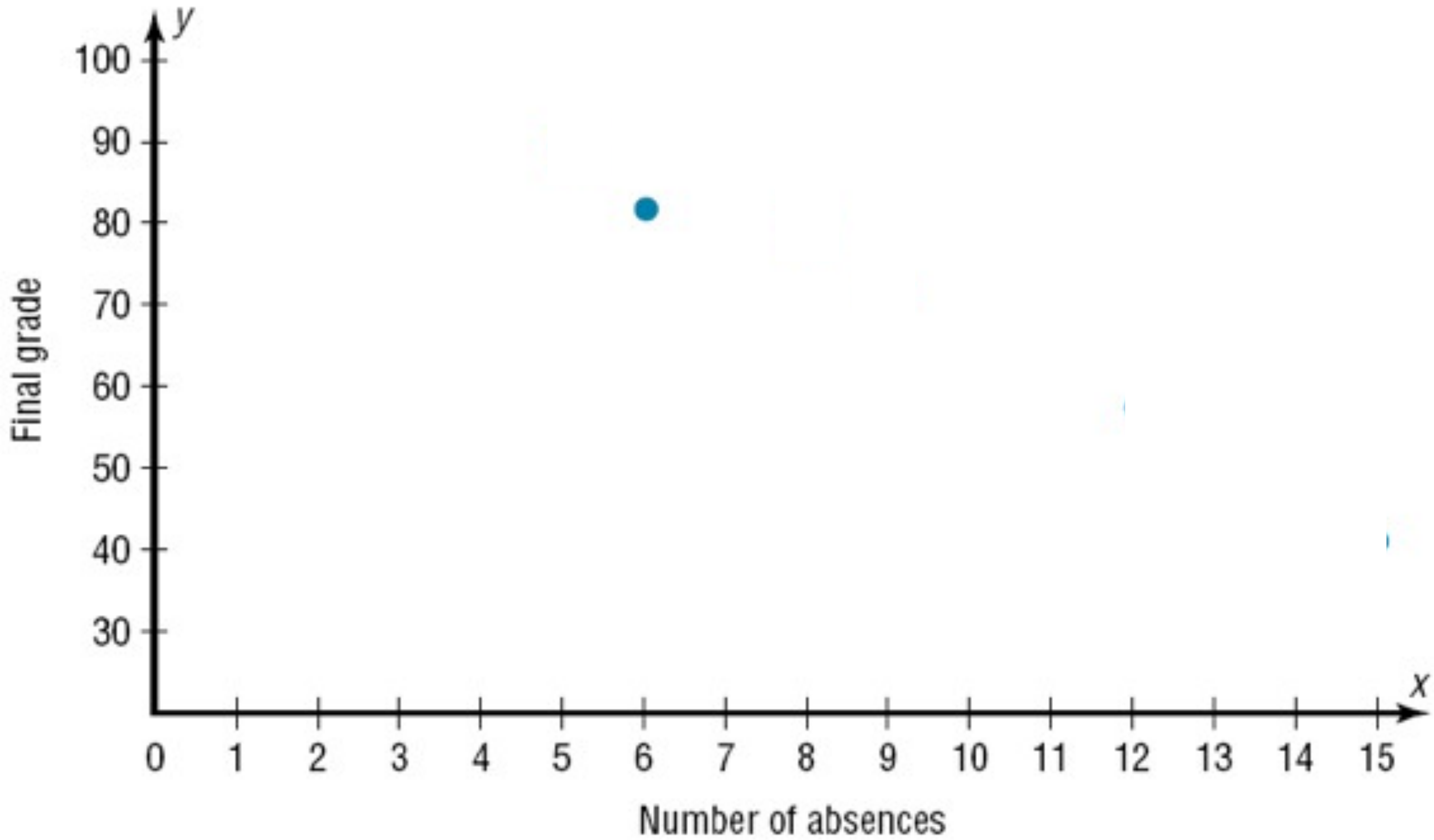
Example 10-2: Absences/Final Grades



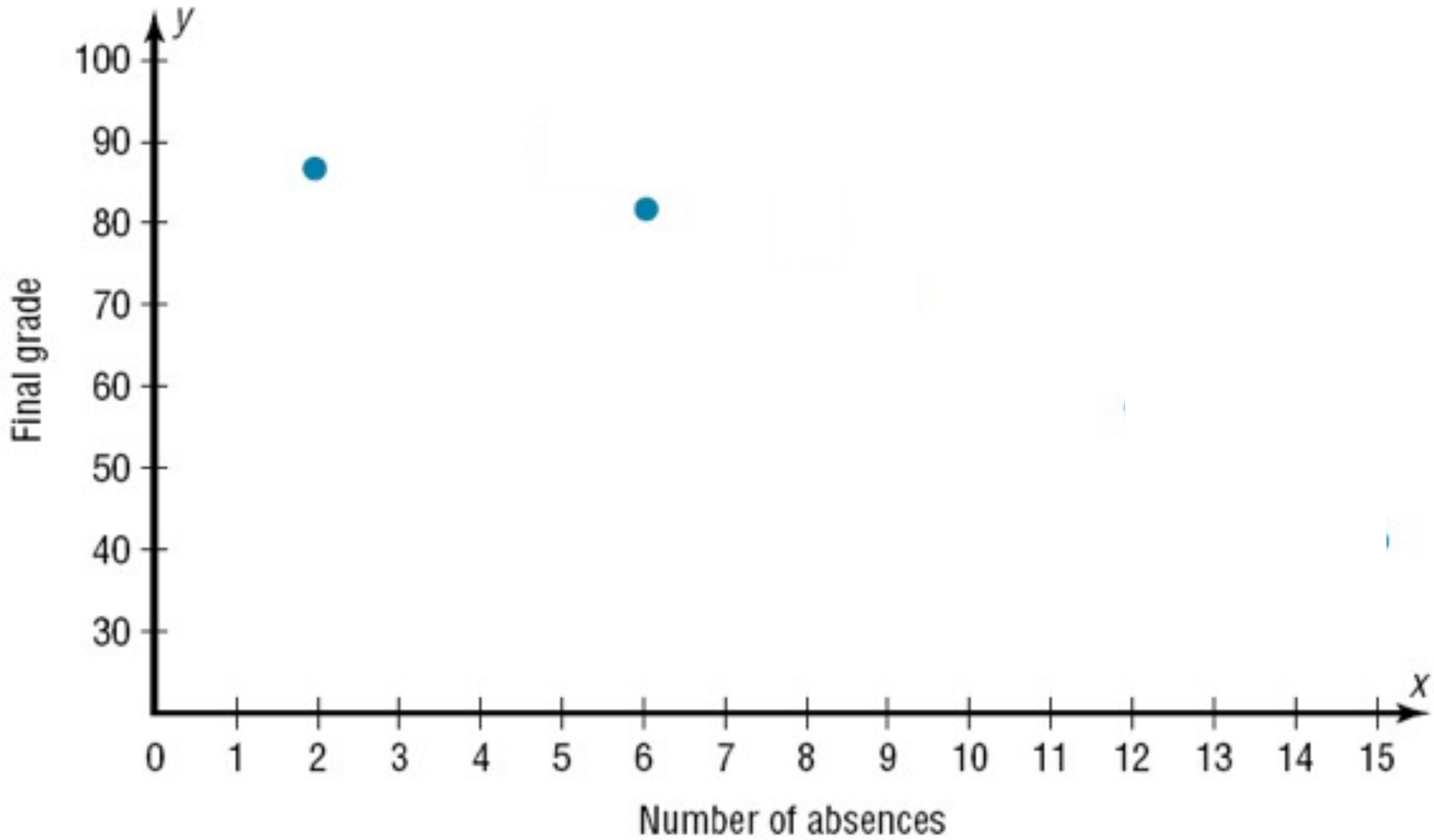
Example 10-2: Absences/Final Grades



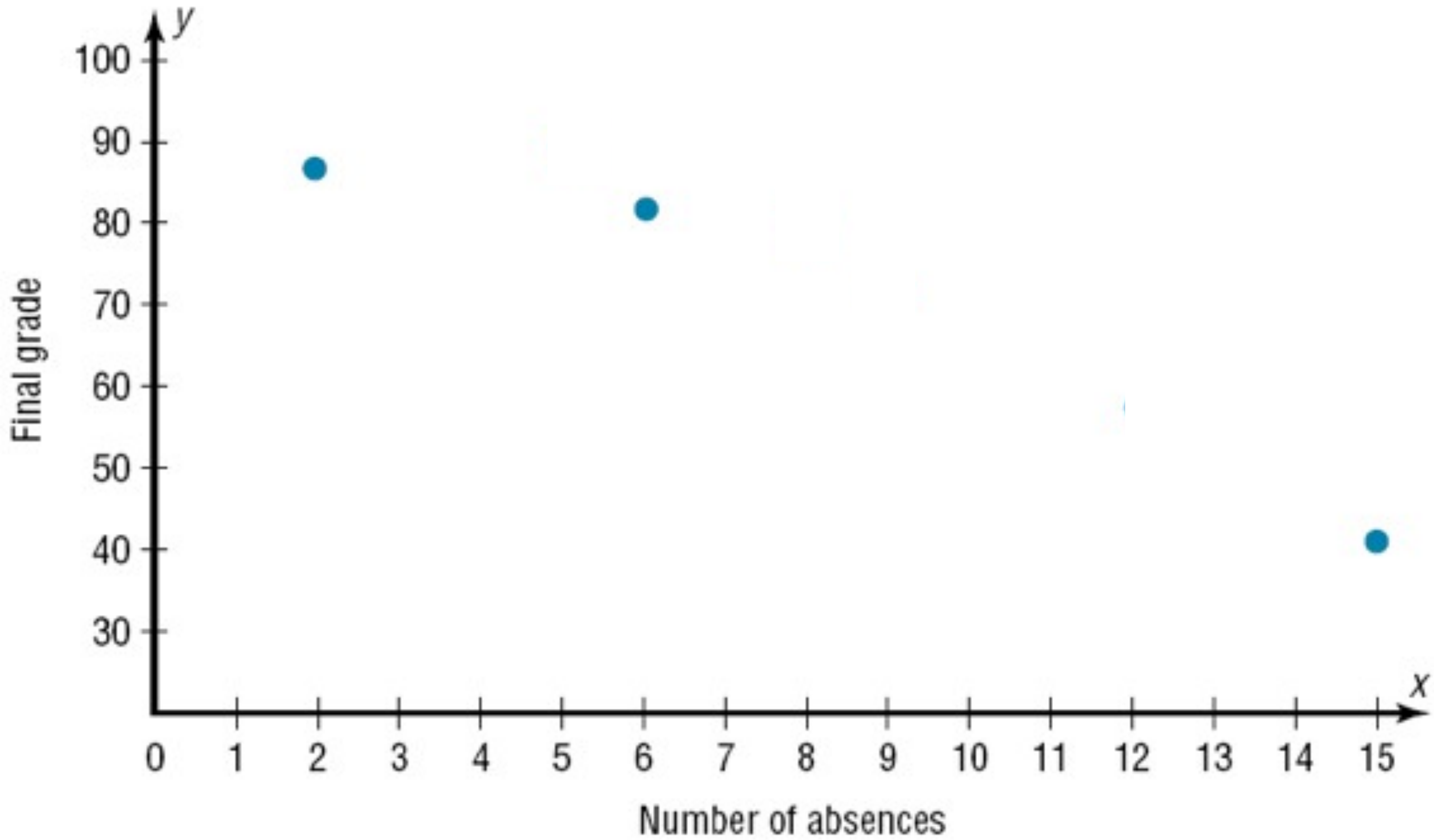
Example 10-2: Absences/Final Grades



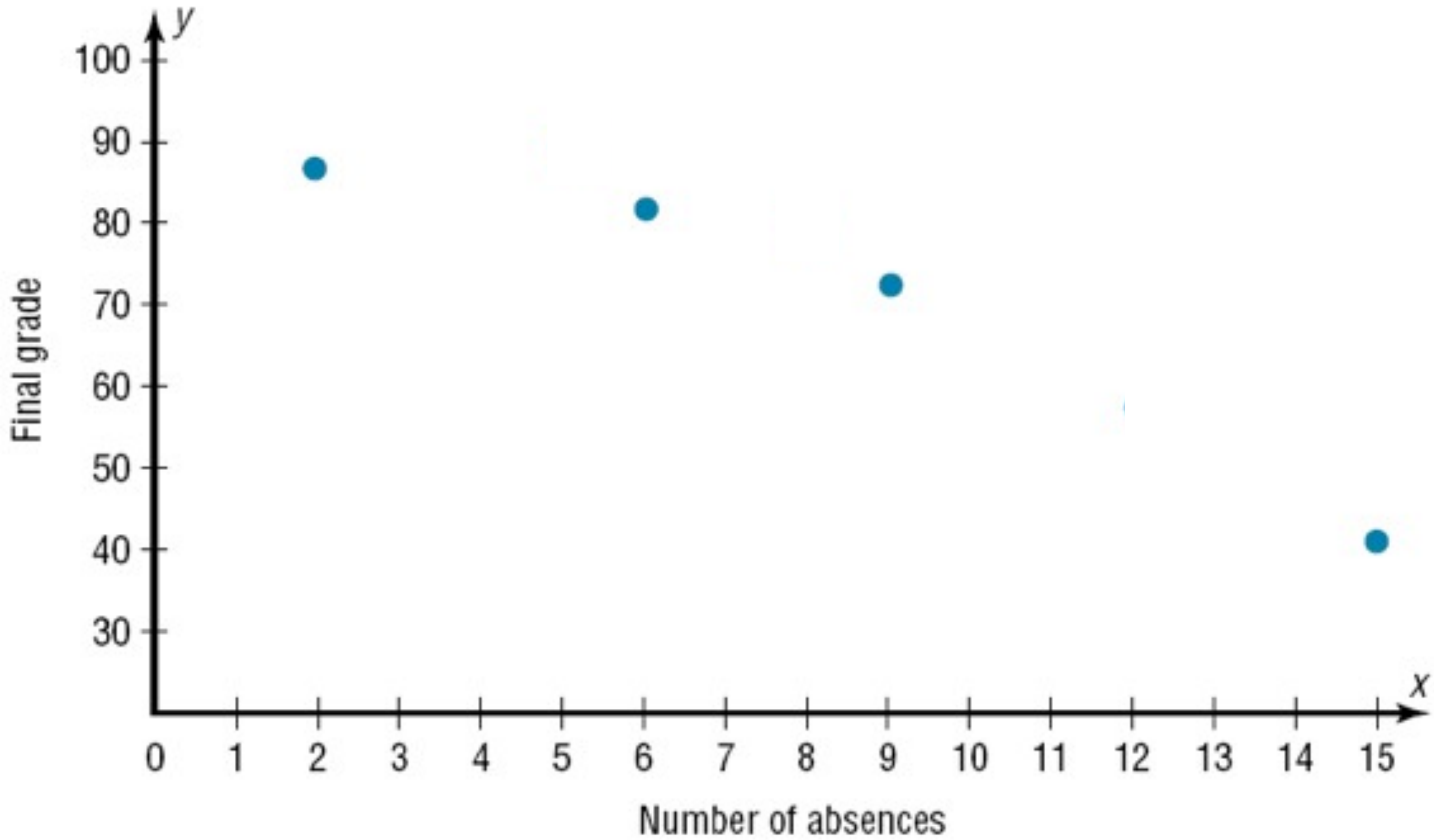
Example 10-2: Absences/Final Grades



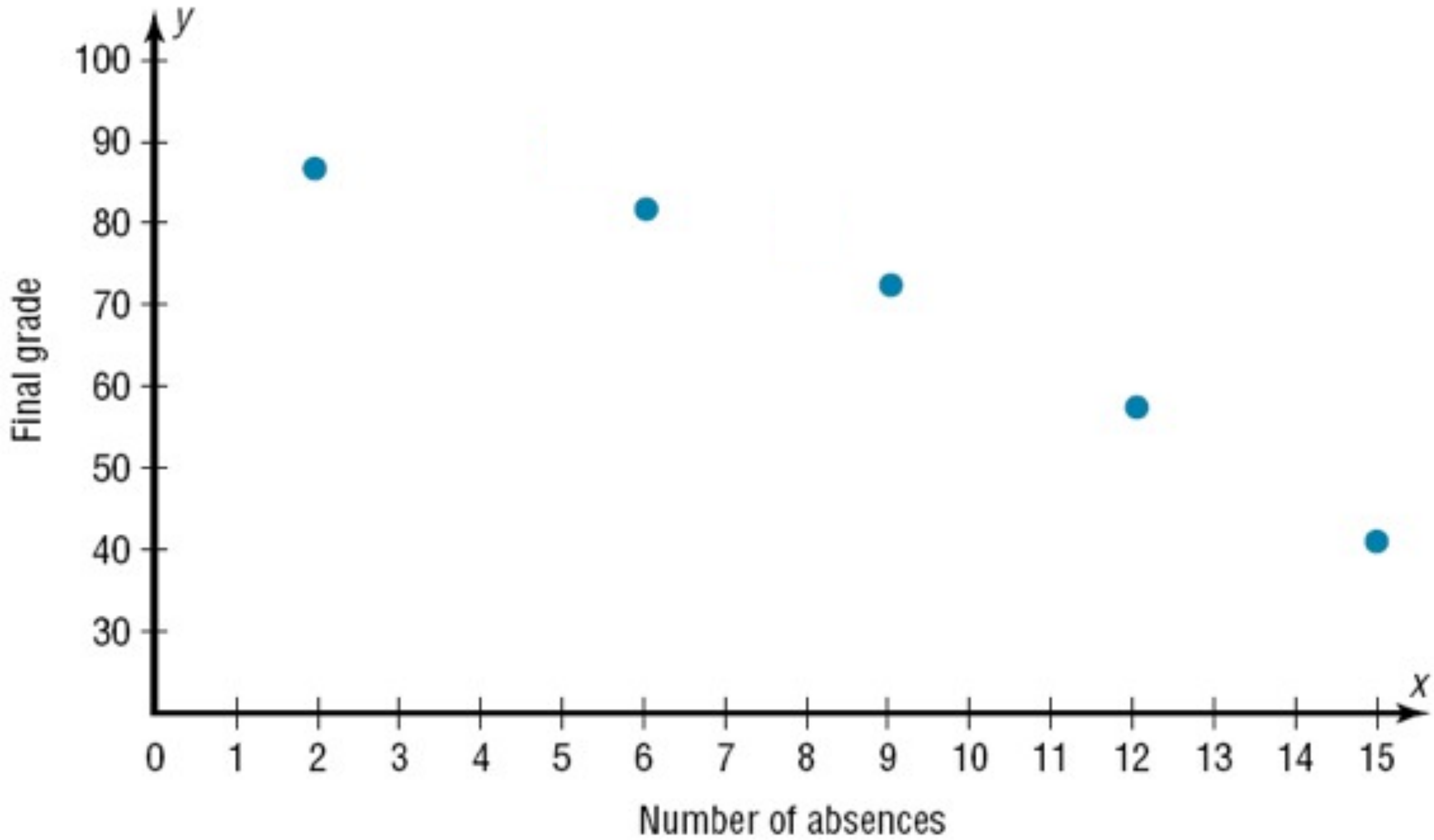
Example 10-2: Absences/Final Grades



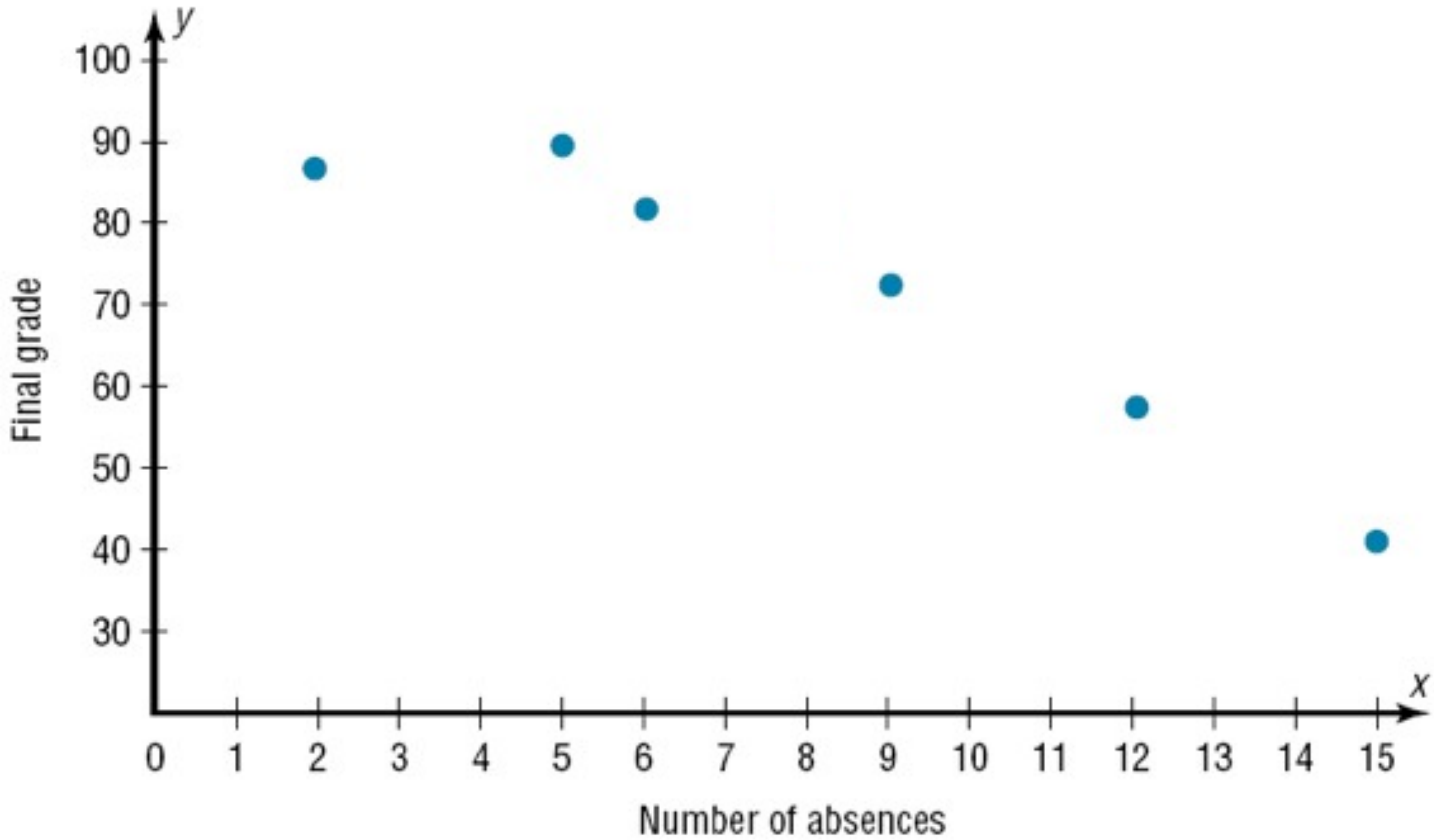
Example 10-2: Absences/Final Grades



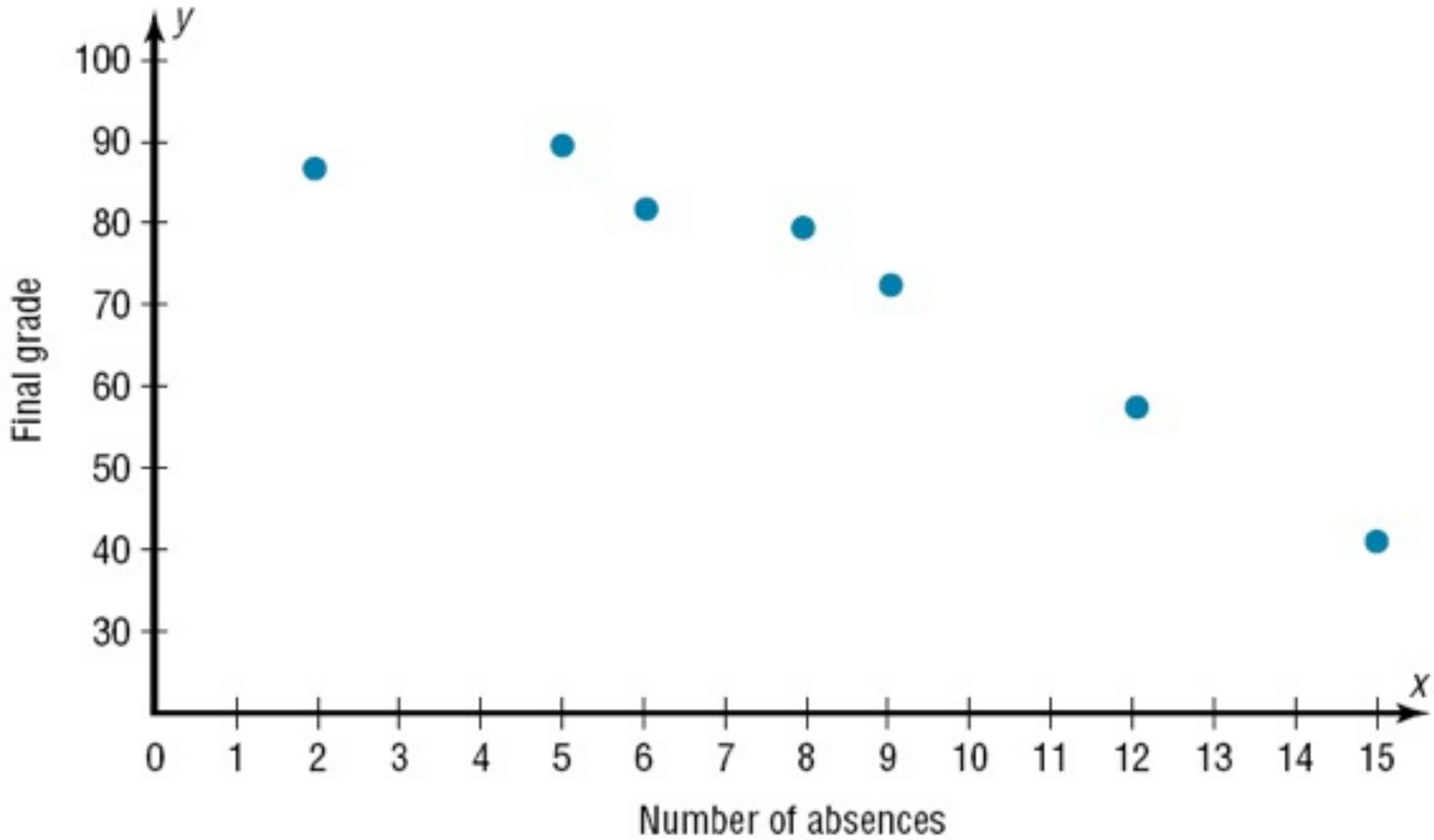
Example 10-2: Absences/Final Grades



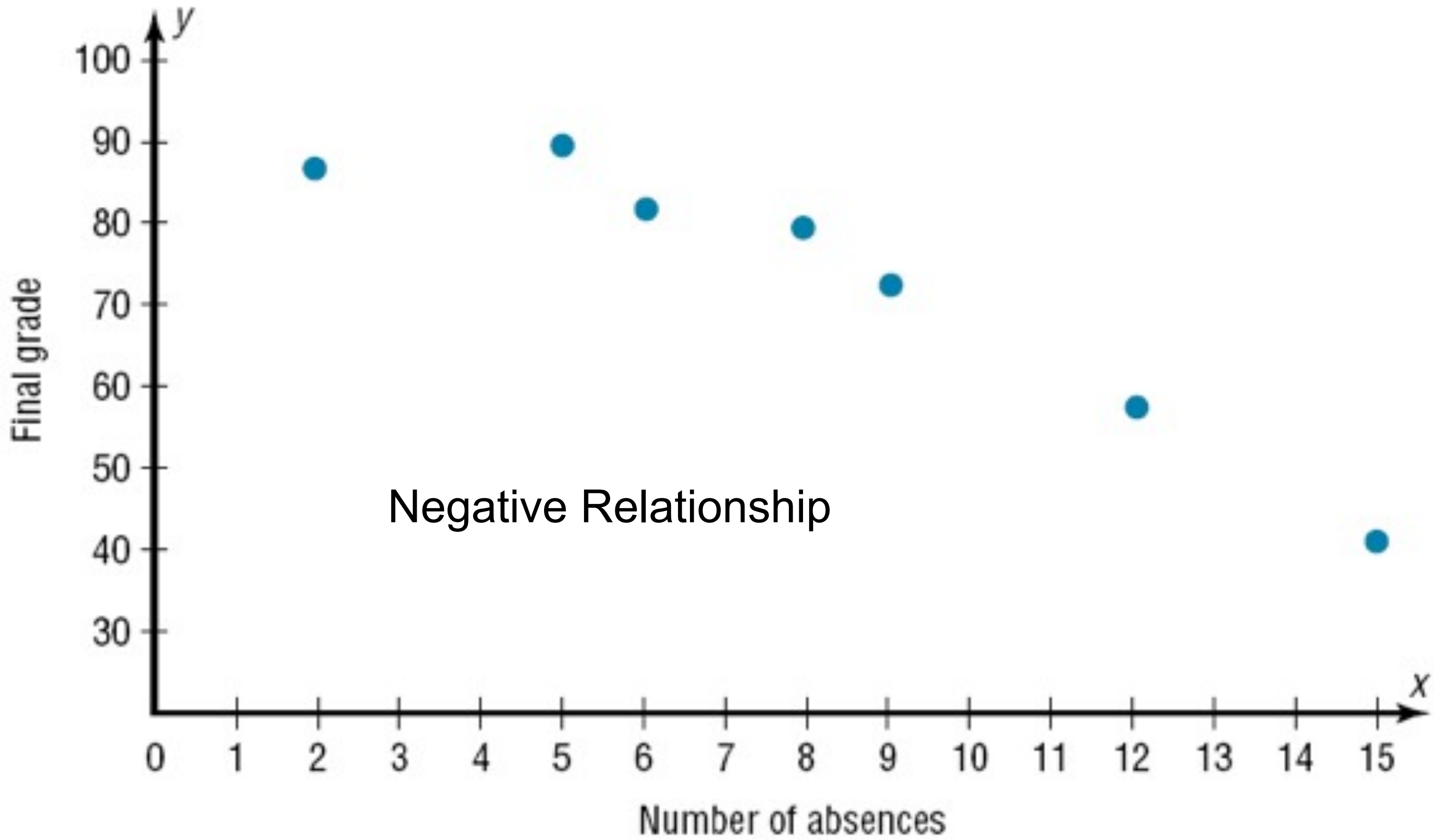
Example 10-2: Absences/Final Grades



Example 10-2: Absences/Final Grades



Example 10-2: Absences/Final Grades



Chapter 10
Correlation and Regression

Section 10-1

Example 10-3

Page #538

Example 10-3: Exercise/Milk Intake

Construct a scatter plot for the data obtained in a study on the number of hours that nine people exercise each week and the amount of milk (in ounces) each person consumes per week.

Subject	Hours x	Amount y
A	3	48
B	0	8
C	2	32
D	5	64
E	8	10
F	5	32
G	10	56
H	2	72
I	1	48

Example 10-3: Exercise/Milk Intake

Construct a scatter plot for the data obtained in a study on the number of hours that nine people exercise each week and the amount of milk (in ounces) each person consumes per week.

Subject	Hours x	Amount y
A	3	48
B	0	8
C	2	32
D	5	64
E	8	10
F	5	32
G	10	56
H	2	72
I	1	48

Step 1: Draw and label the x and y axes.

Example 10-3: Exercise/Milk Intake

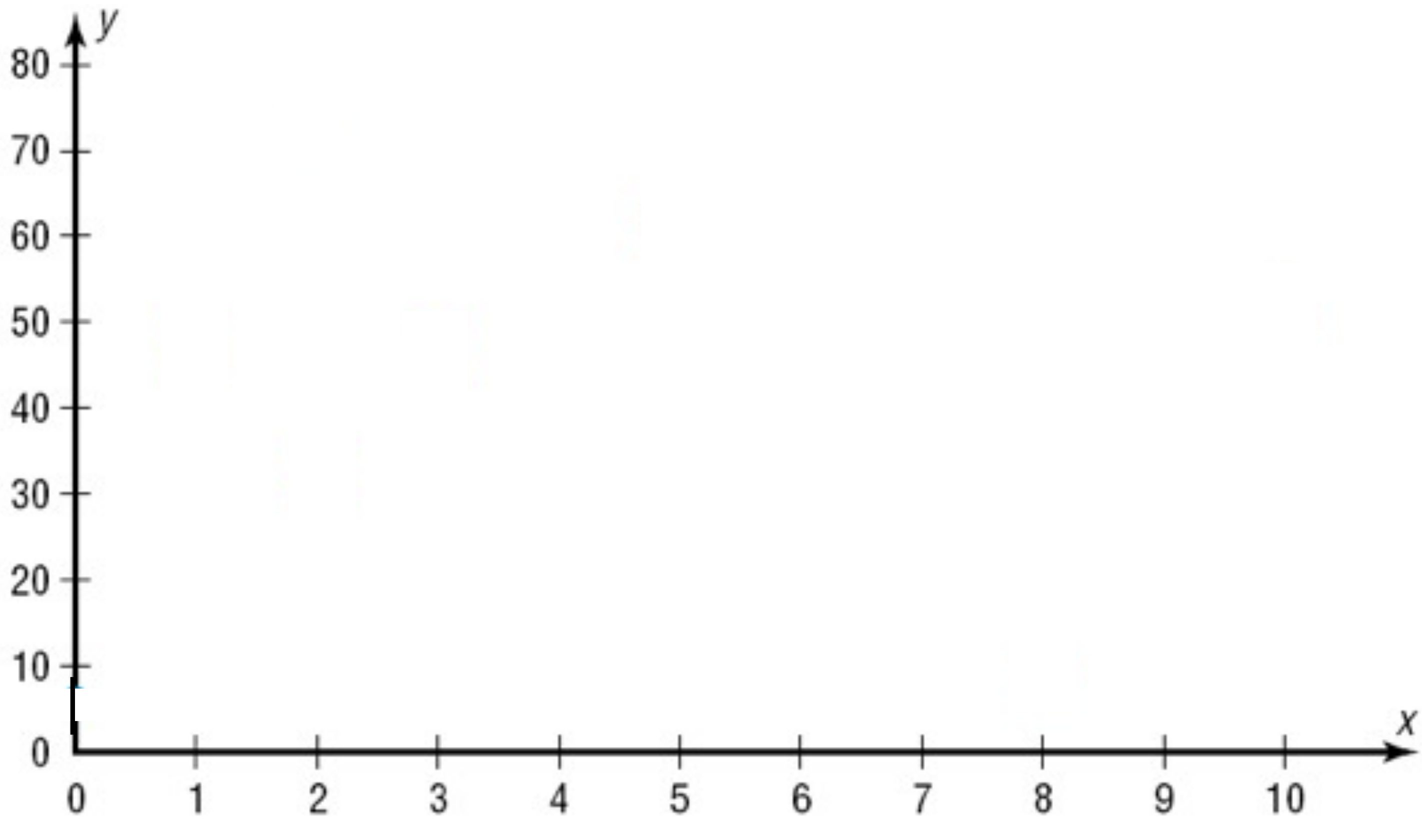
Construct a scatter plot for the data obtained in a study on the number of hours that nine people exercise each week and the amount of milk (in ounces) each person consumes per week.

Subject	Hours x	Amount y
A	3	48
B	0	8
C	2	32
D	5	64
E	8	10
F	5	32
G	10	56
H	2	72
I	1	48

Step 1: Draw and label the *x* and *y* axes.

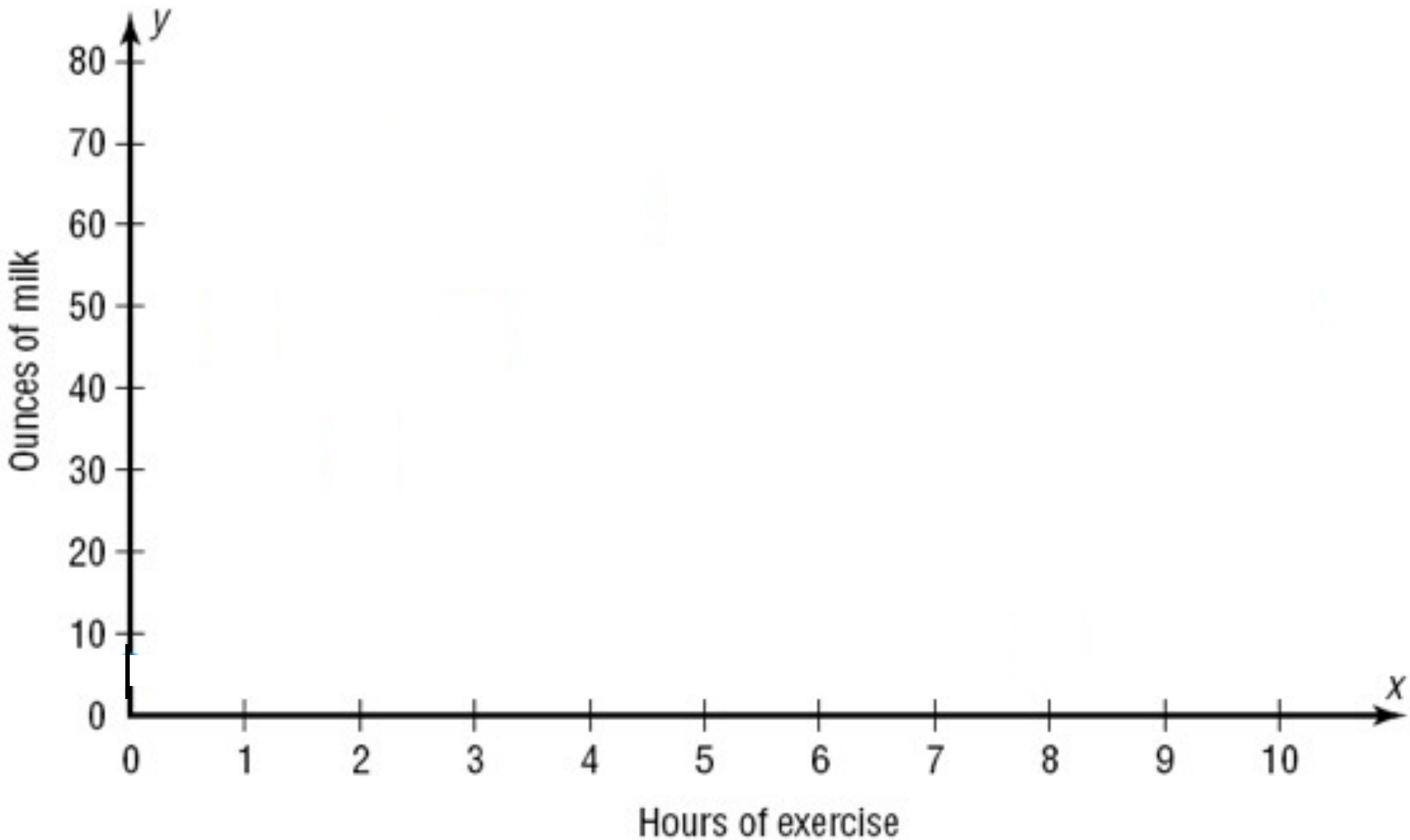
Step 2: Plot each point on the graph.

Example 10-3: Exercise/Milk Intake

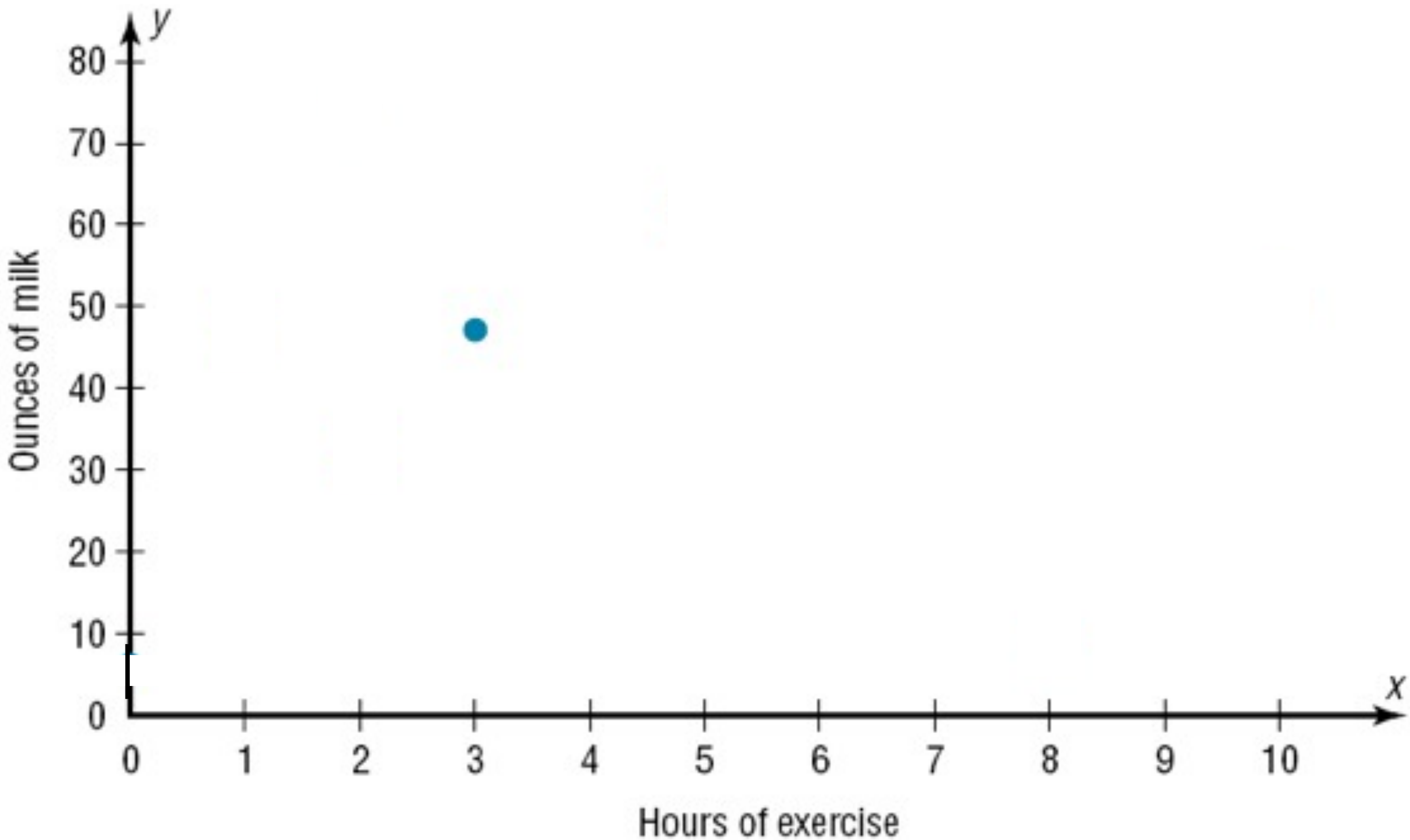


Example 10-3: Exercise/Milk Intake

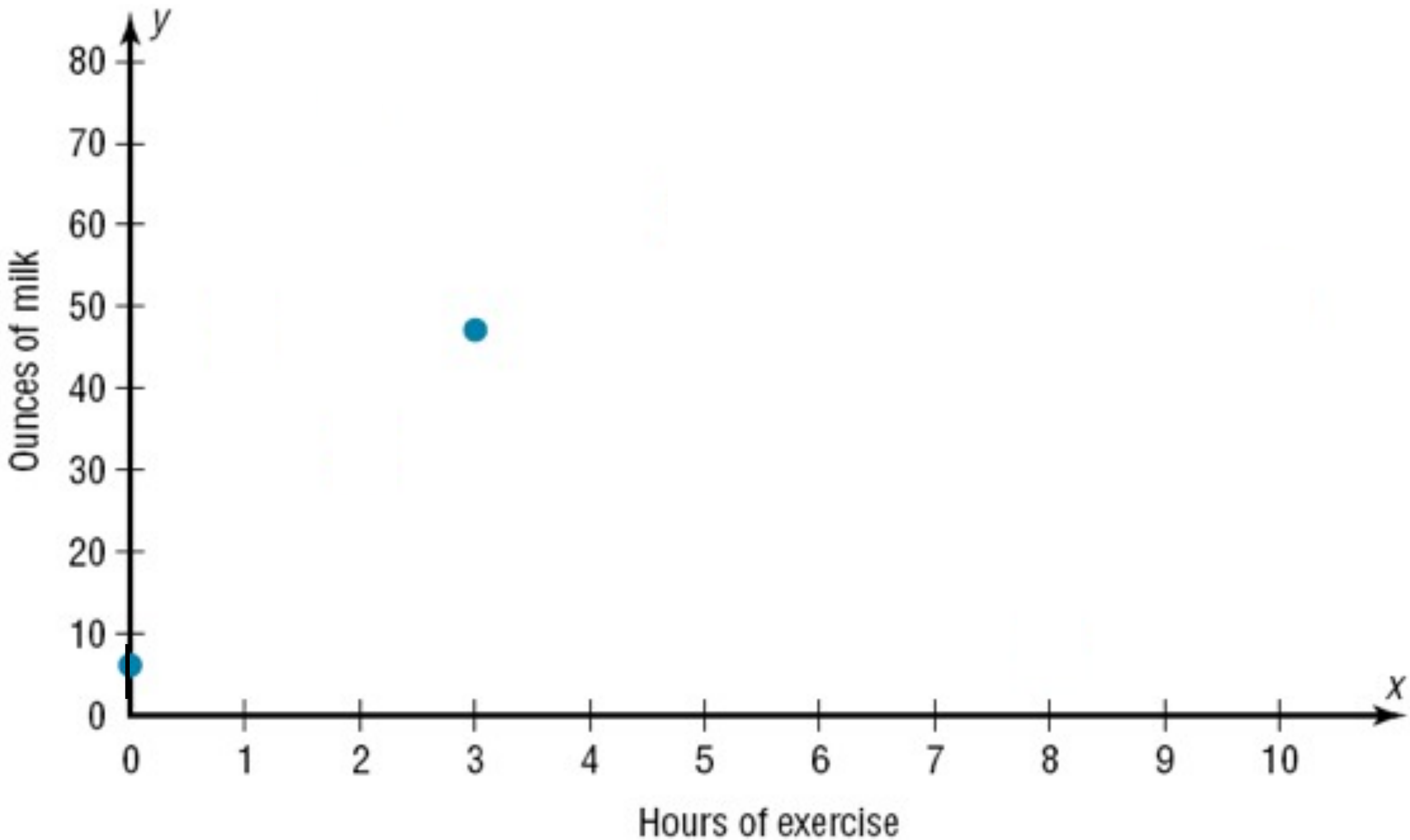
Example 10-3: Exercise/Milk Intake



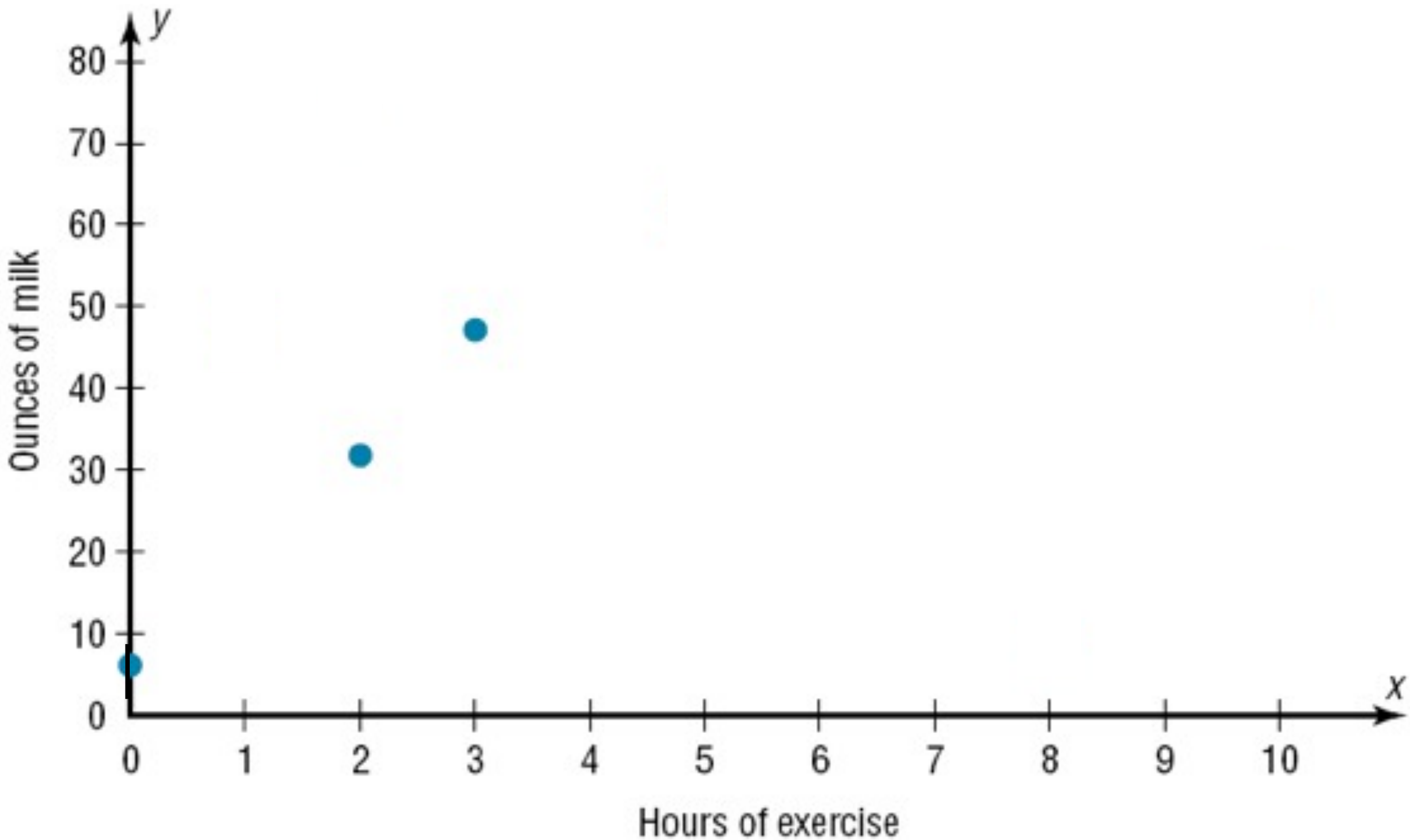
Example 10-3: Exercise/Milk Intake



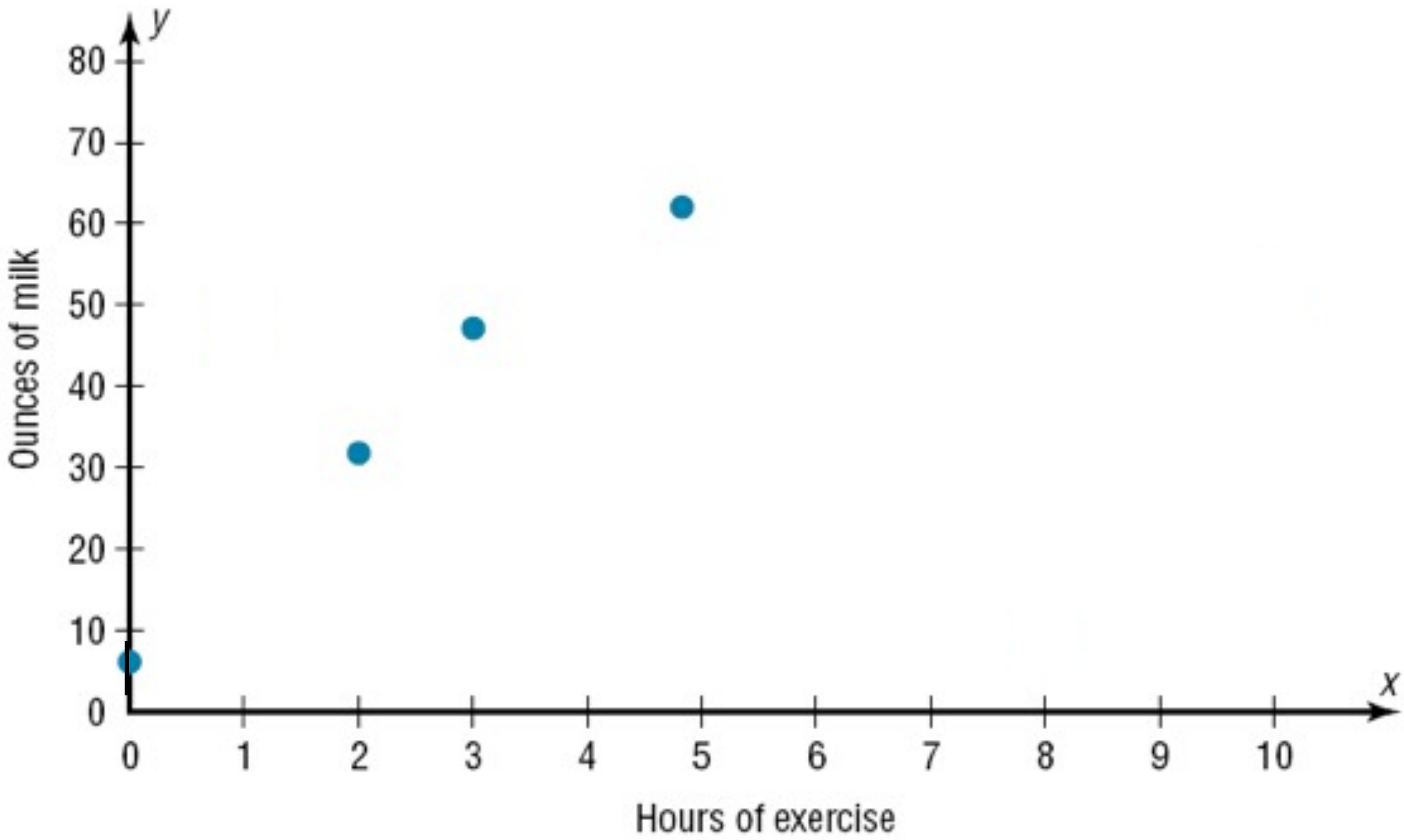
Example 10-3: Exercise/Milk Intake



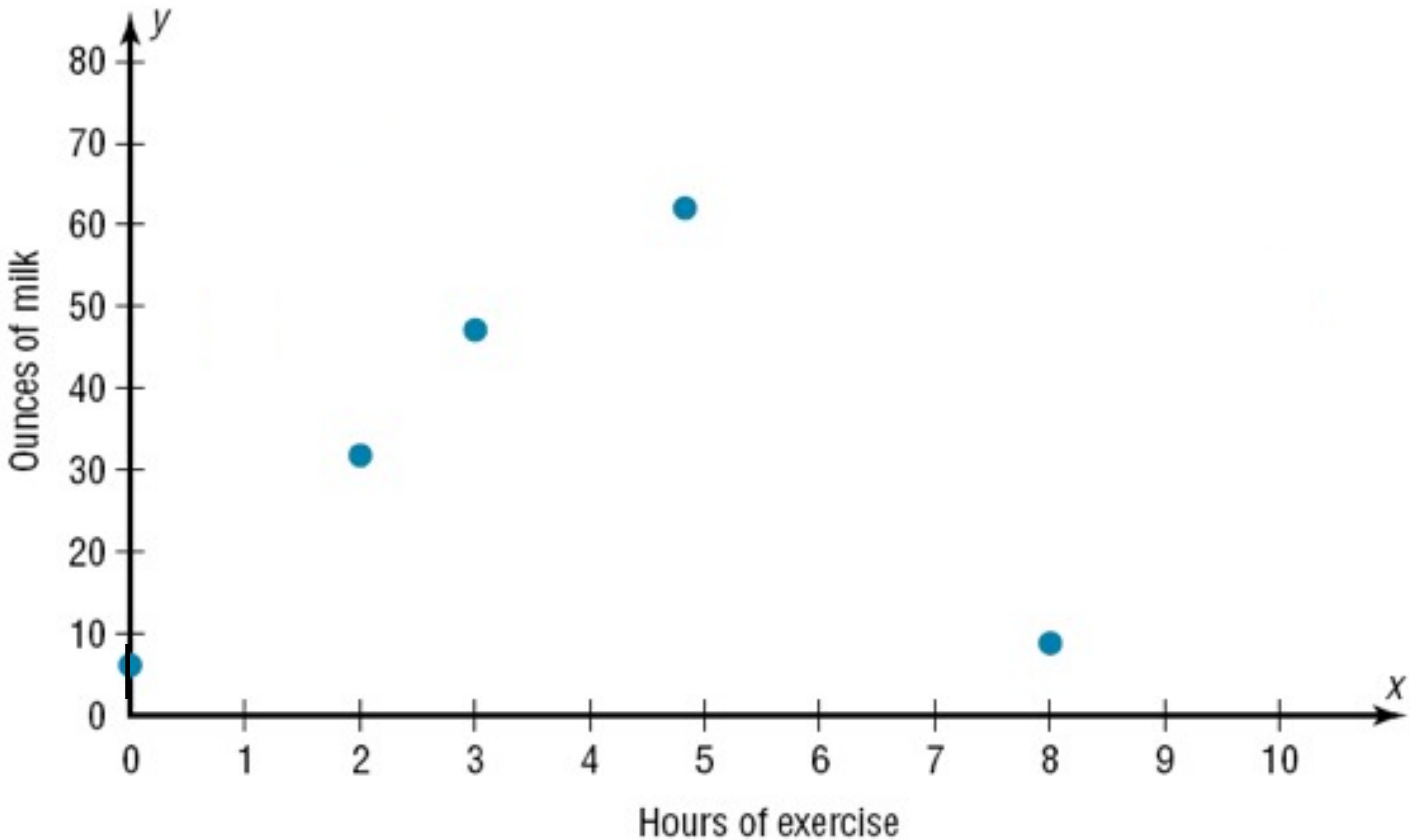
Example 10-3: Exercise/Milk Intake



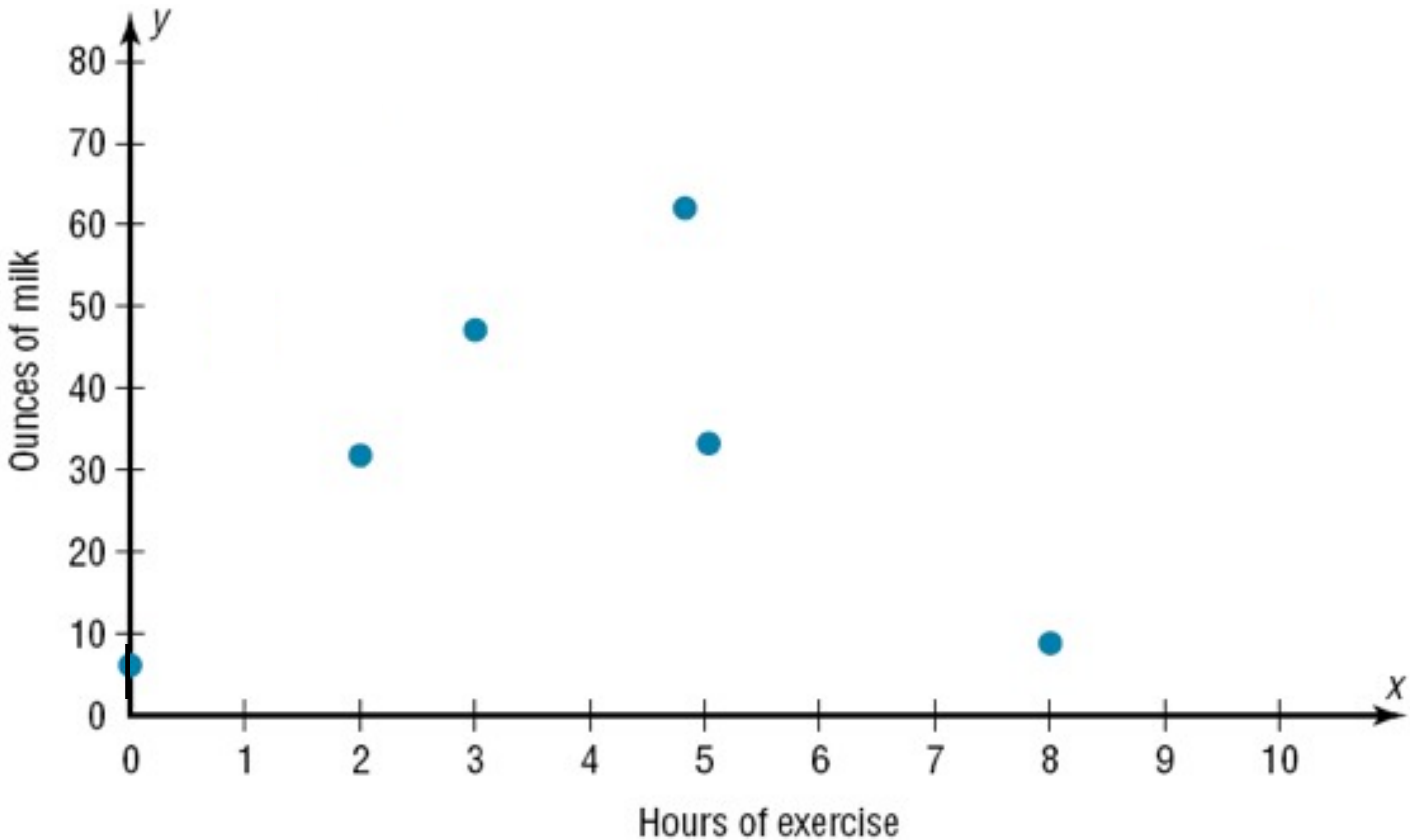
Example 10-3: Exercise/Milk Intake



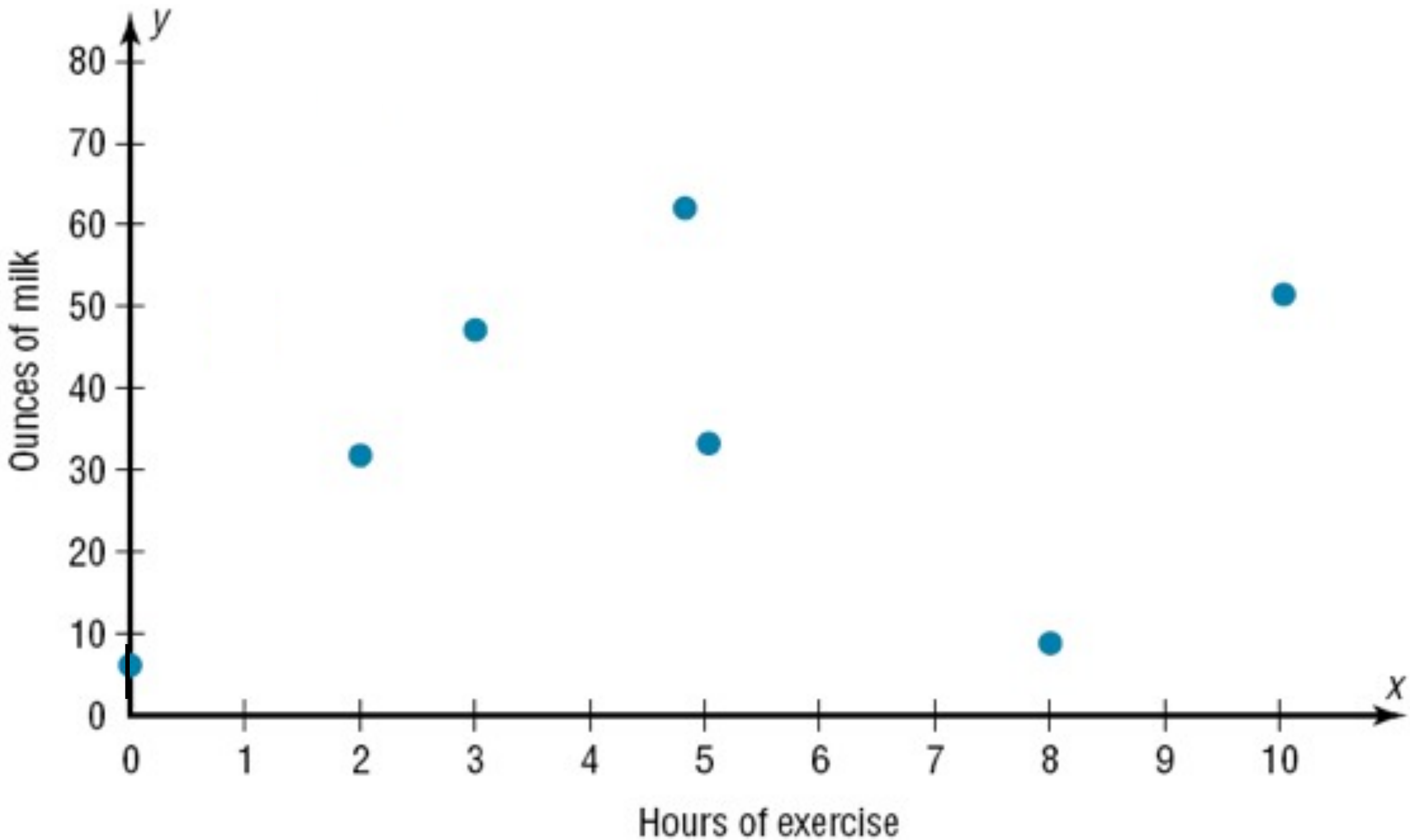
Example 10-3: Exercise/Milk Intake



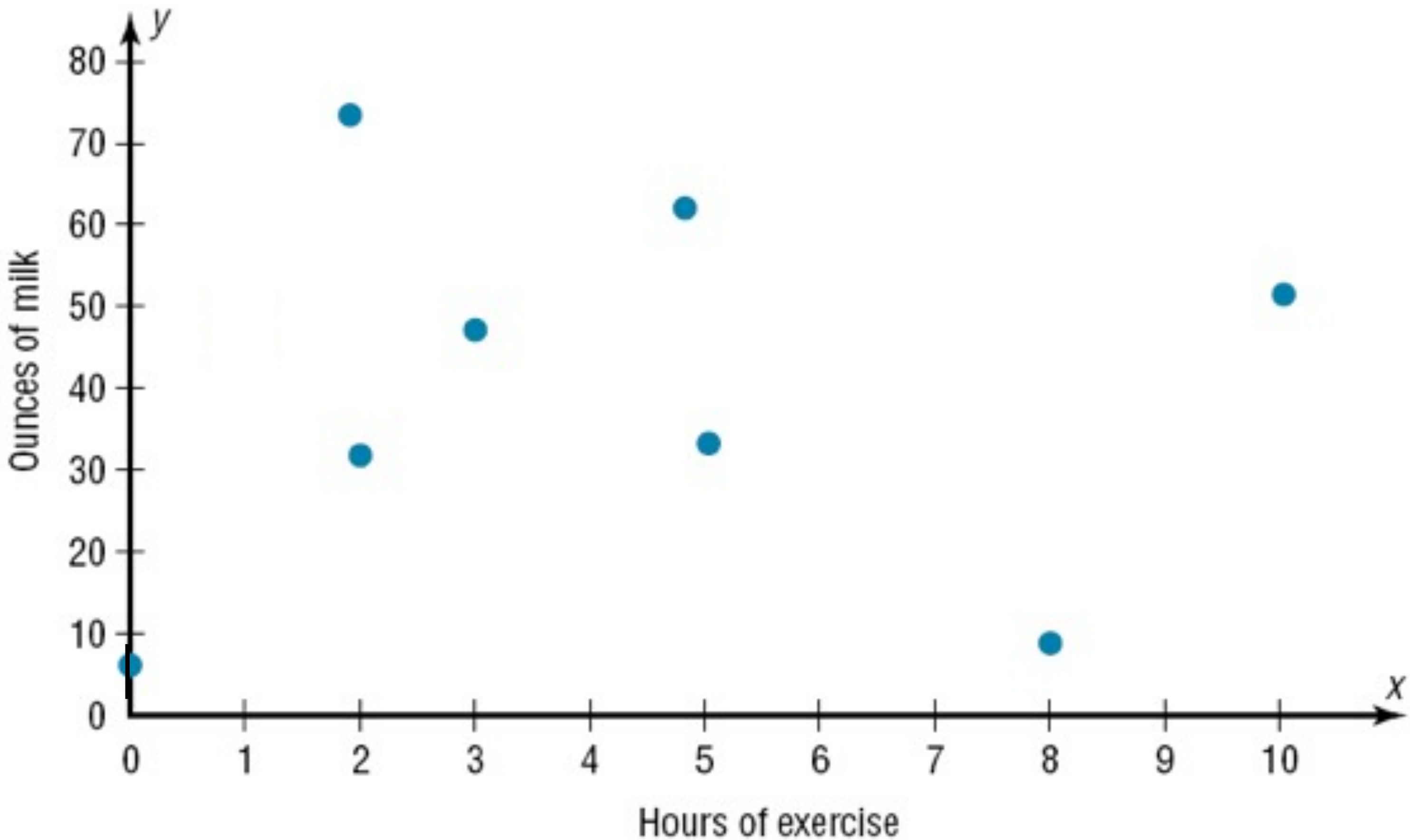
Example 10-3: Exercise/Milk Intake



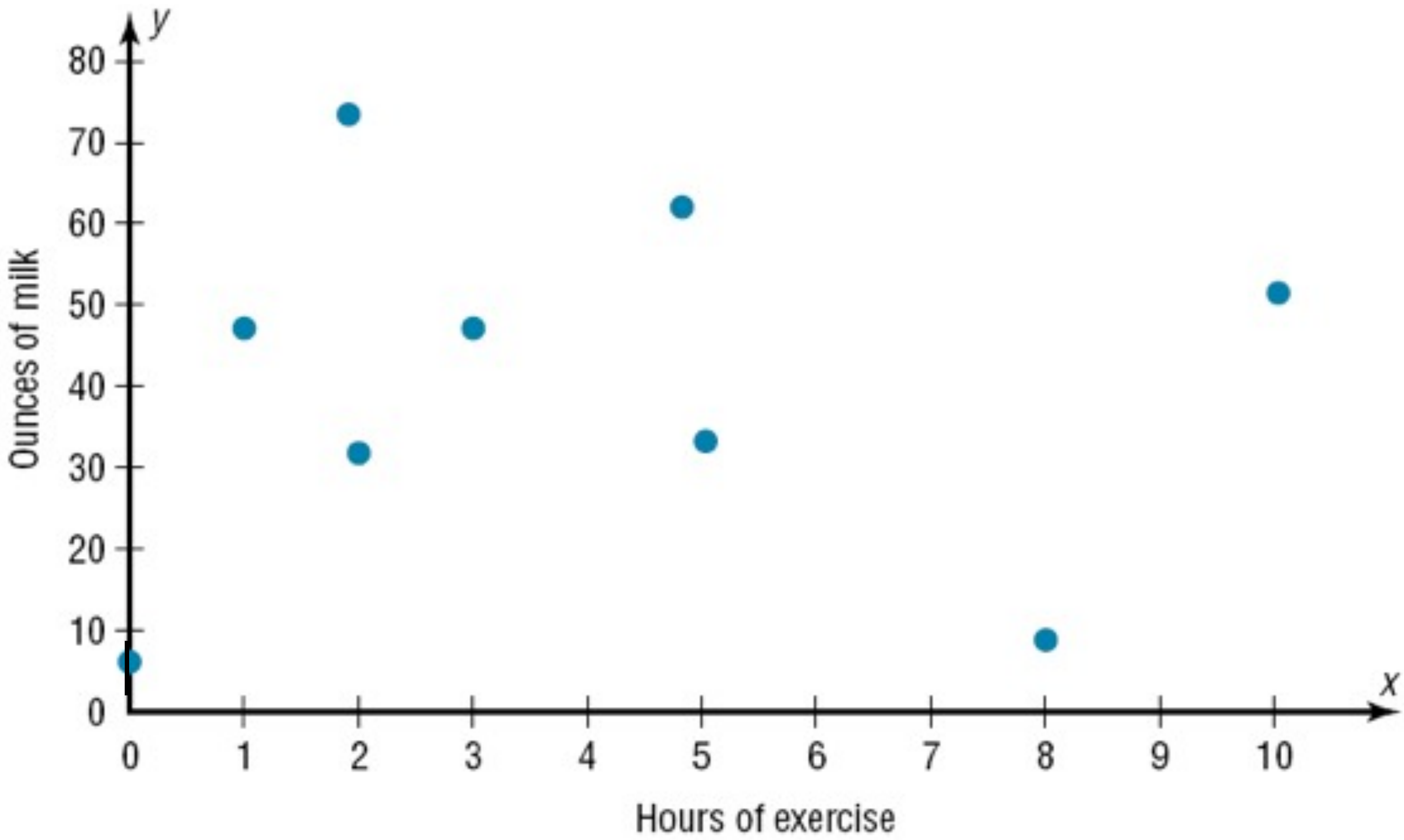
Example 10-3: Exercise/Milk Intake



Example 10-3: Exercise/Milk Intake



Example 10-3: Exercise/Milk Intake



Example 10-3: Exercise/Milk Intake

Correlation

- The **correlation coefficient** computed from the sample data measures the strength and direction of a linear relationship between two variables.

Correlation

- The **correlation coefficient** computed from the sample data measures the strength and direction of a linear relationship between two variables.
- There are several types of correlation coefficients. The one explained in this section is called the **Pearson product moment correlation coefficient (PPMC)**.

Correlation

- The **correlation coefficient** computed from the sample data measures the strength and direction of a linear relationship between two variables.
- There are several types of correlation coefficients. The one explained in this section is called the **Pearson product moment correlation coefficient (PPMC)**.
- The symbol for the sample correlation coefficient is r. The symbol for the population correlation coefficient is ρ.

Correlation

- The range of the correlation coefficient is from -1 to $+1$.

Correlation

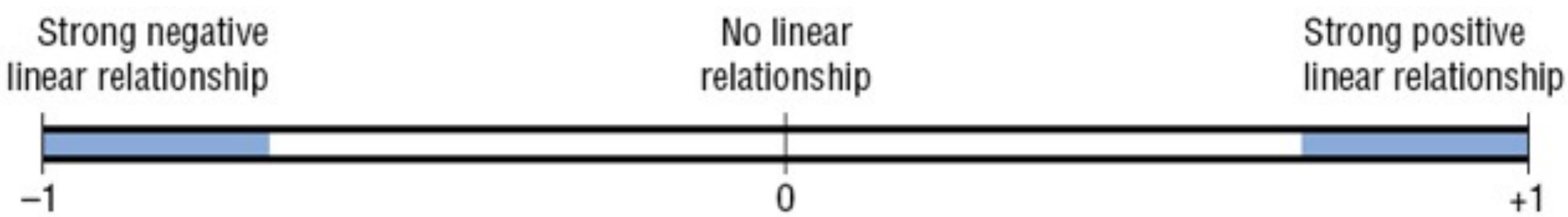
- The range of the correlation coefficient is from –1 to +1.
- If there is a **strong positive linear relationship** between the variables, the value of *r* will be close to +1.

Correlation

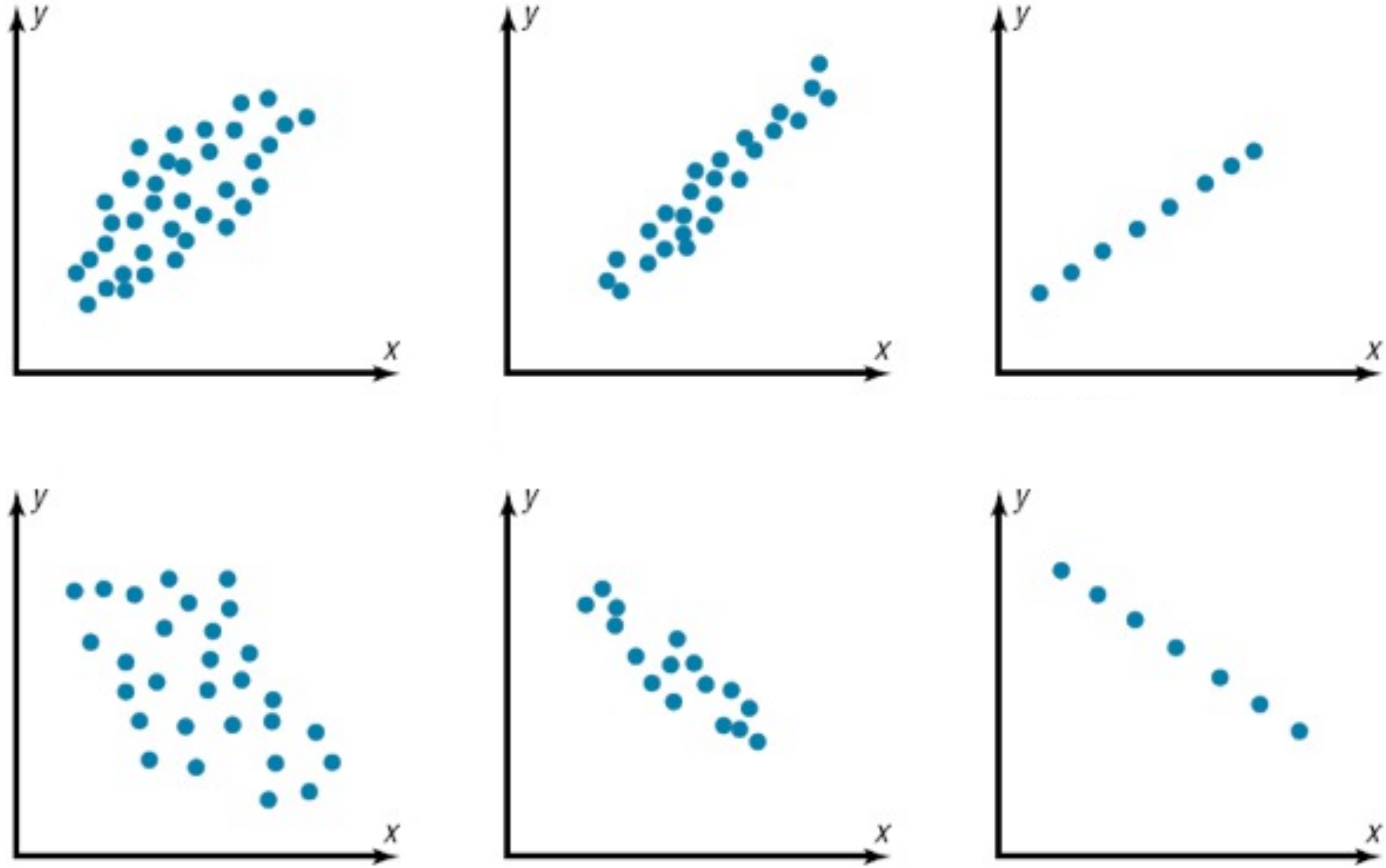
- The range of the correlation coefficient is from –1 to +1.
- If there is a **strong positive linear relationship** between the variables, the value of r will be close to +1.
- If there is a **strong negative linear relationship** between the variables, the value of r will be close to –1.

Correlation

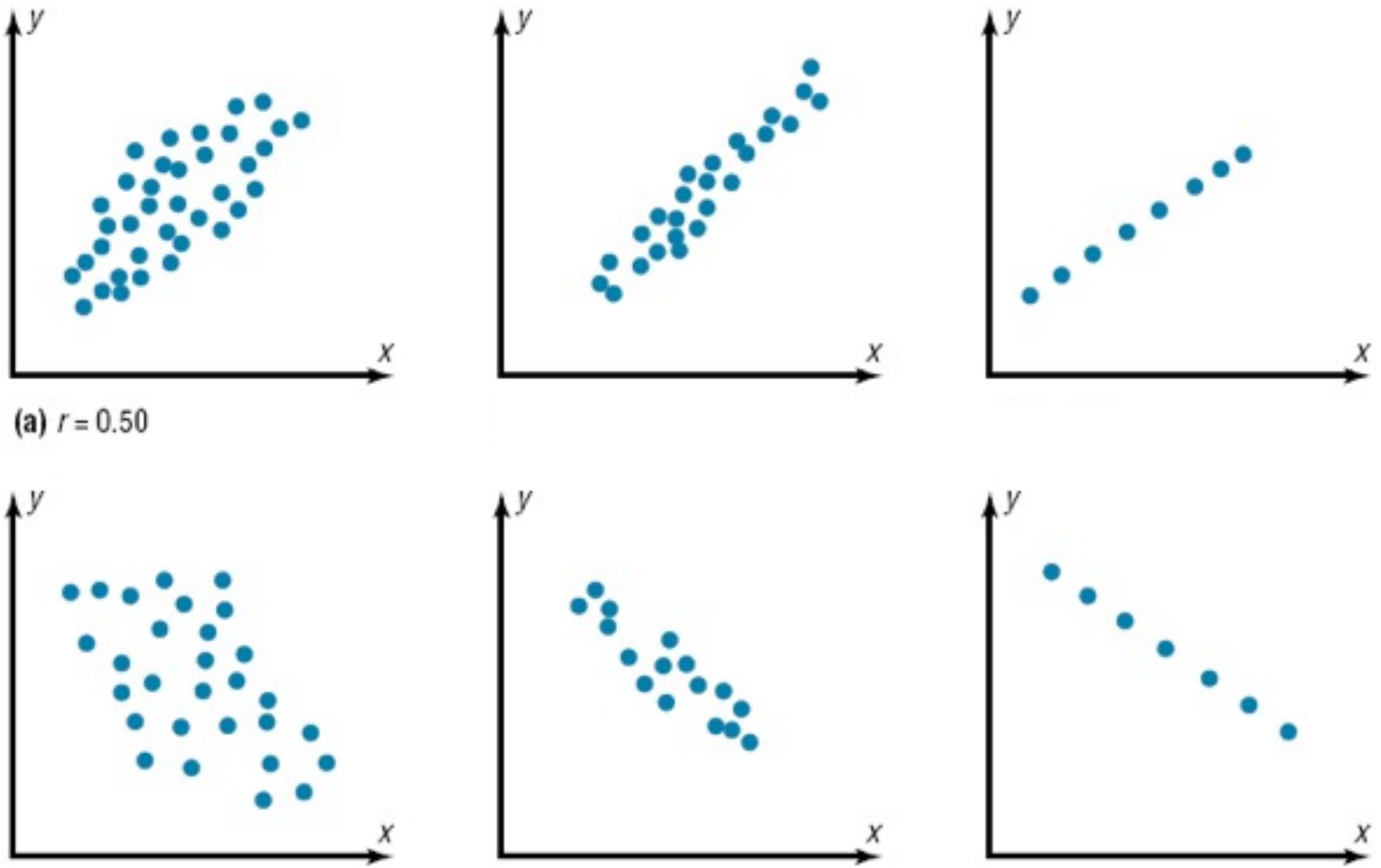
- The range of the correlation coefficient is from –1 to +1.
- If there is a **strong positive linear relationship** between the variables, the value of *r* will be close to +1.
- If there is a **strong negative linear relationship** between the variables, the value of *r* will be close to –1.



Correlation

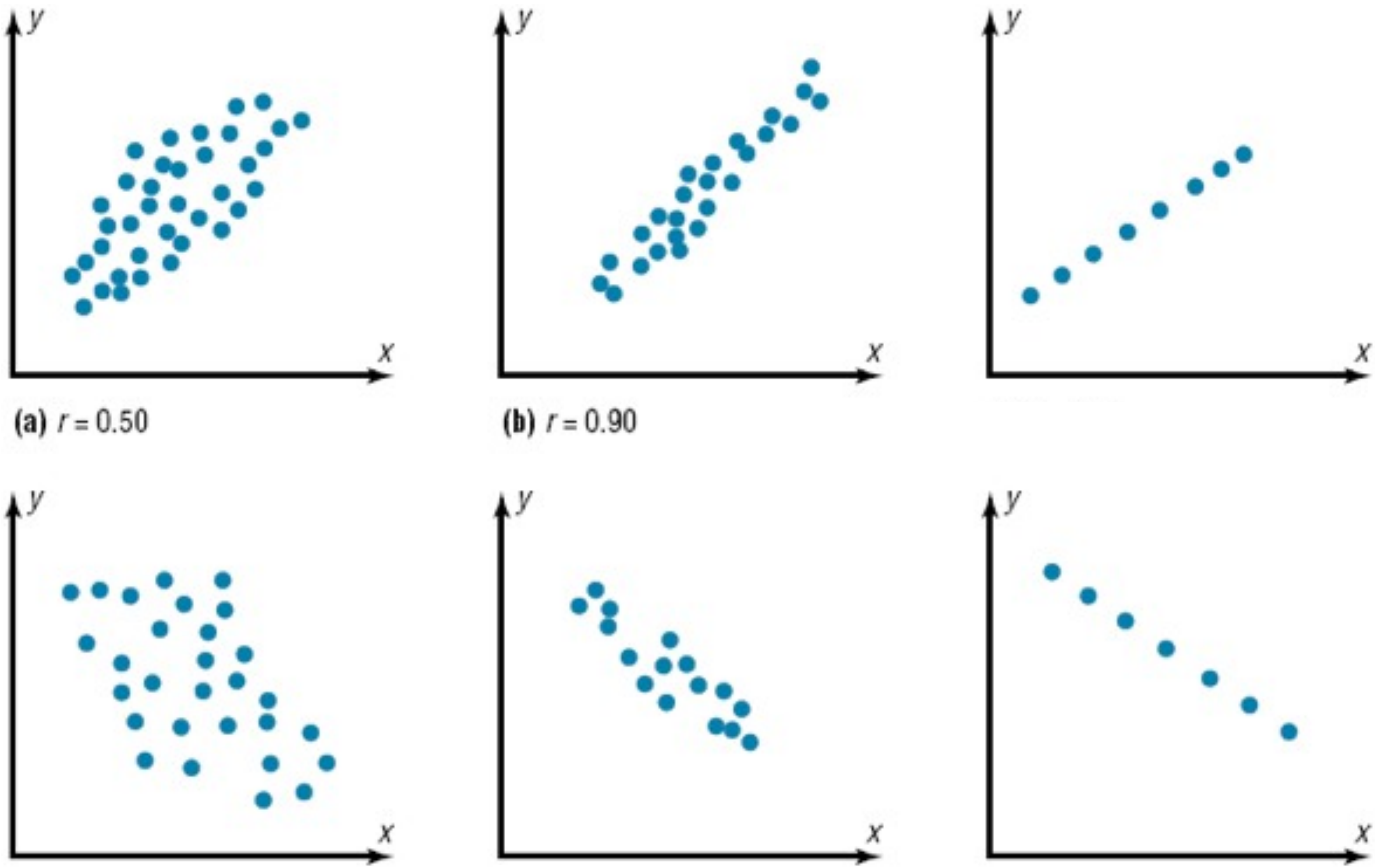


Correlation



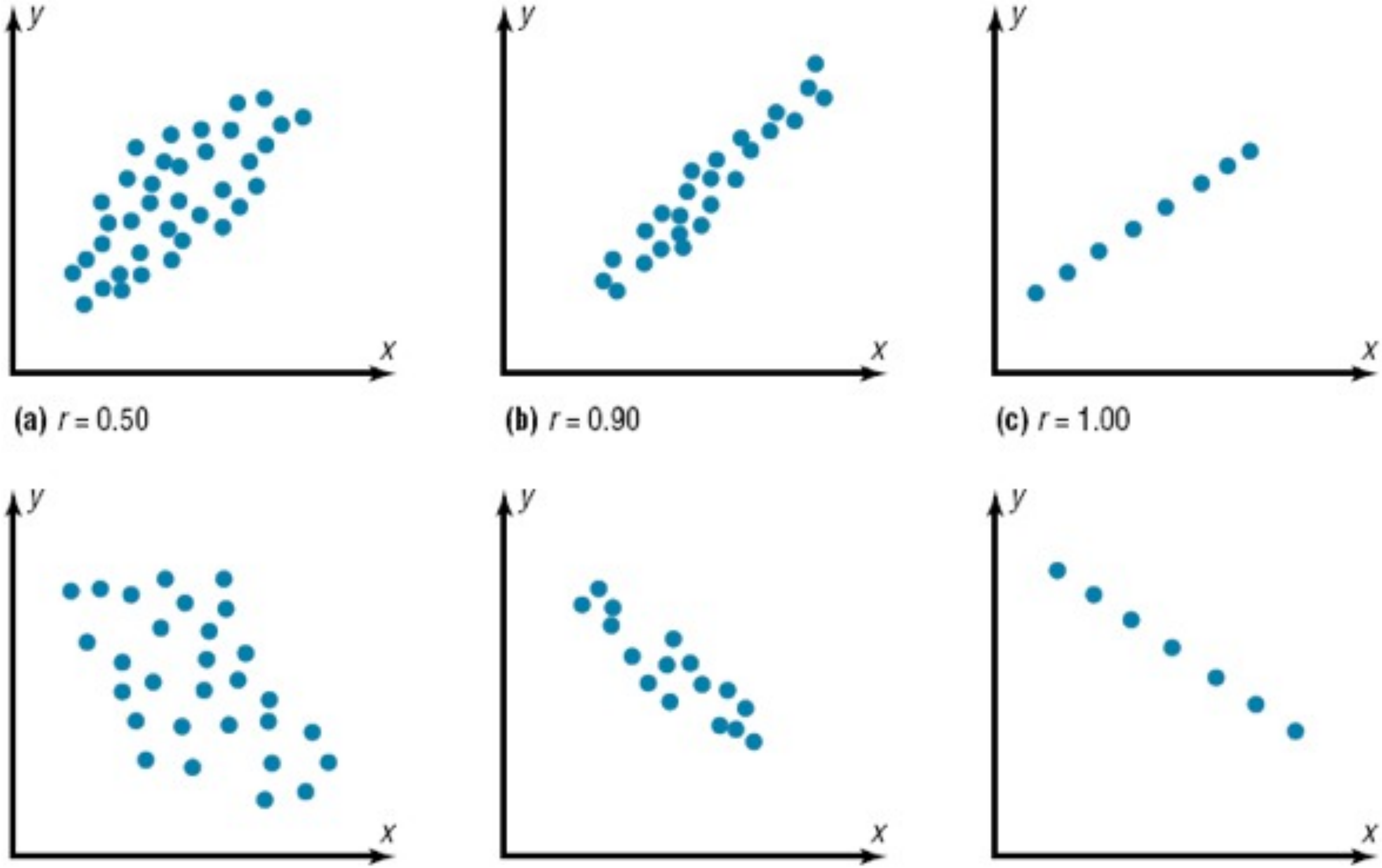
Bluman, Chapter 10

Correlation

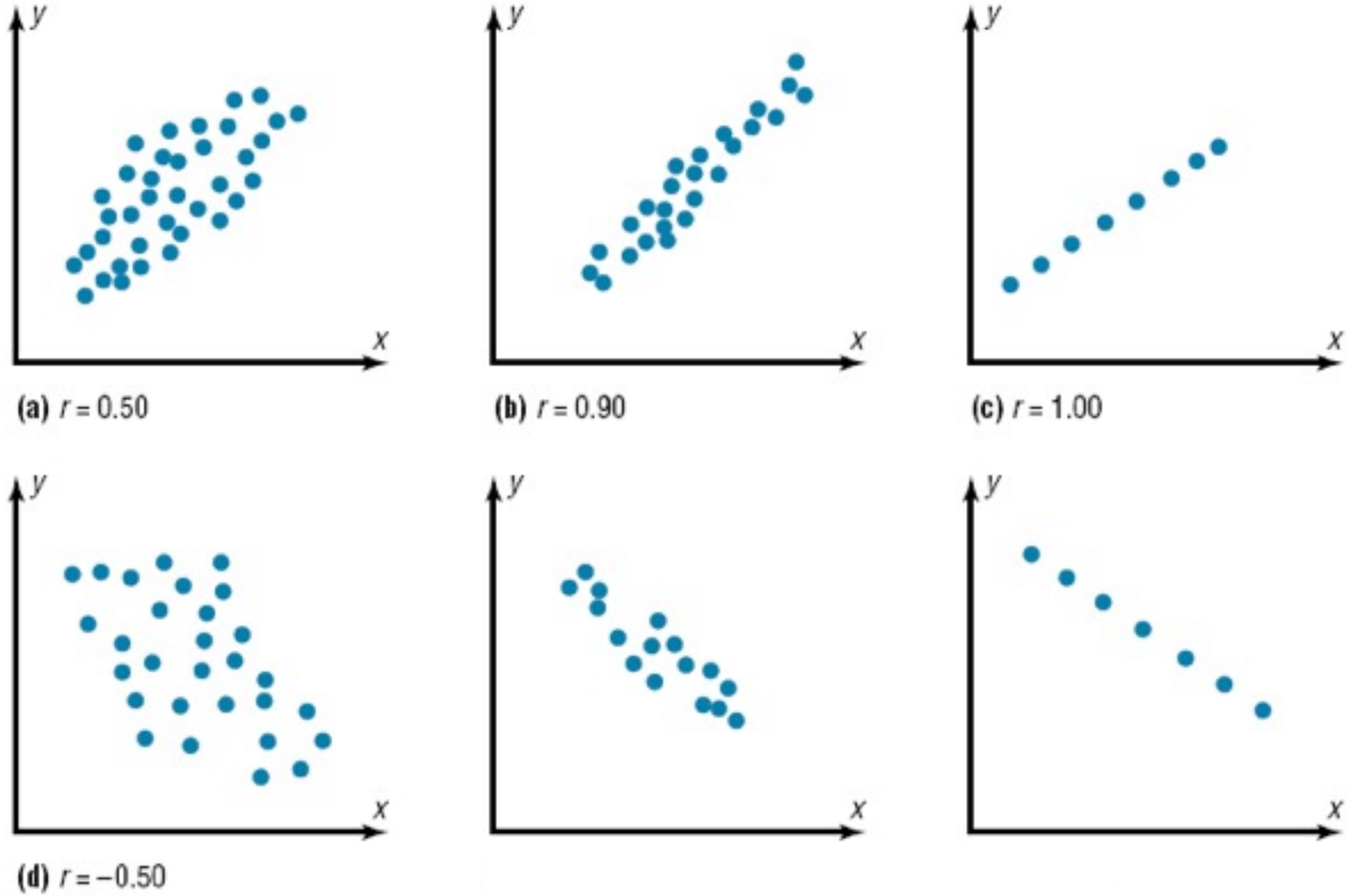


Bluman, Chapter 10

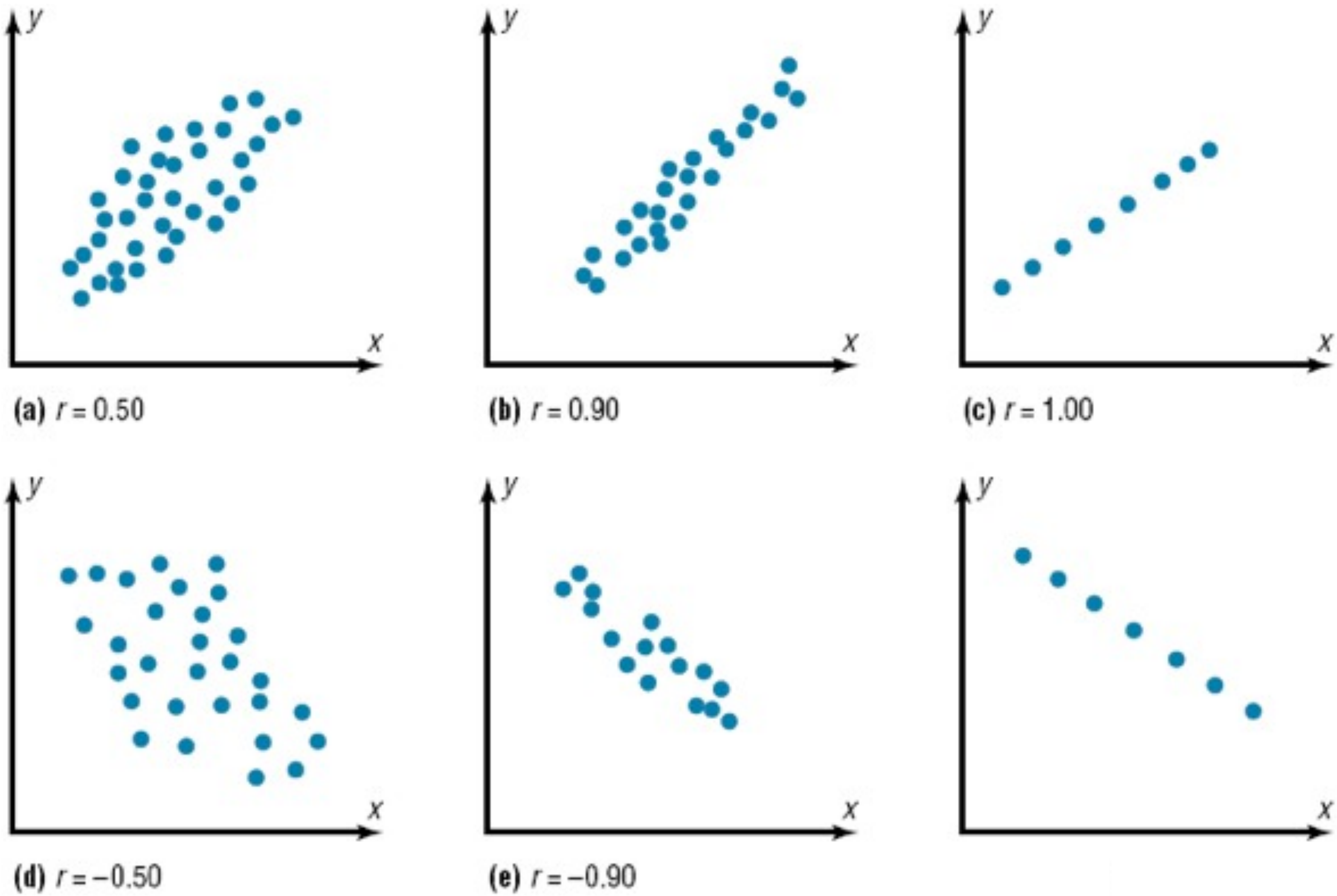
Correlation



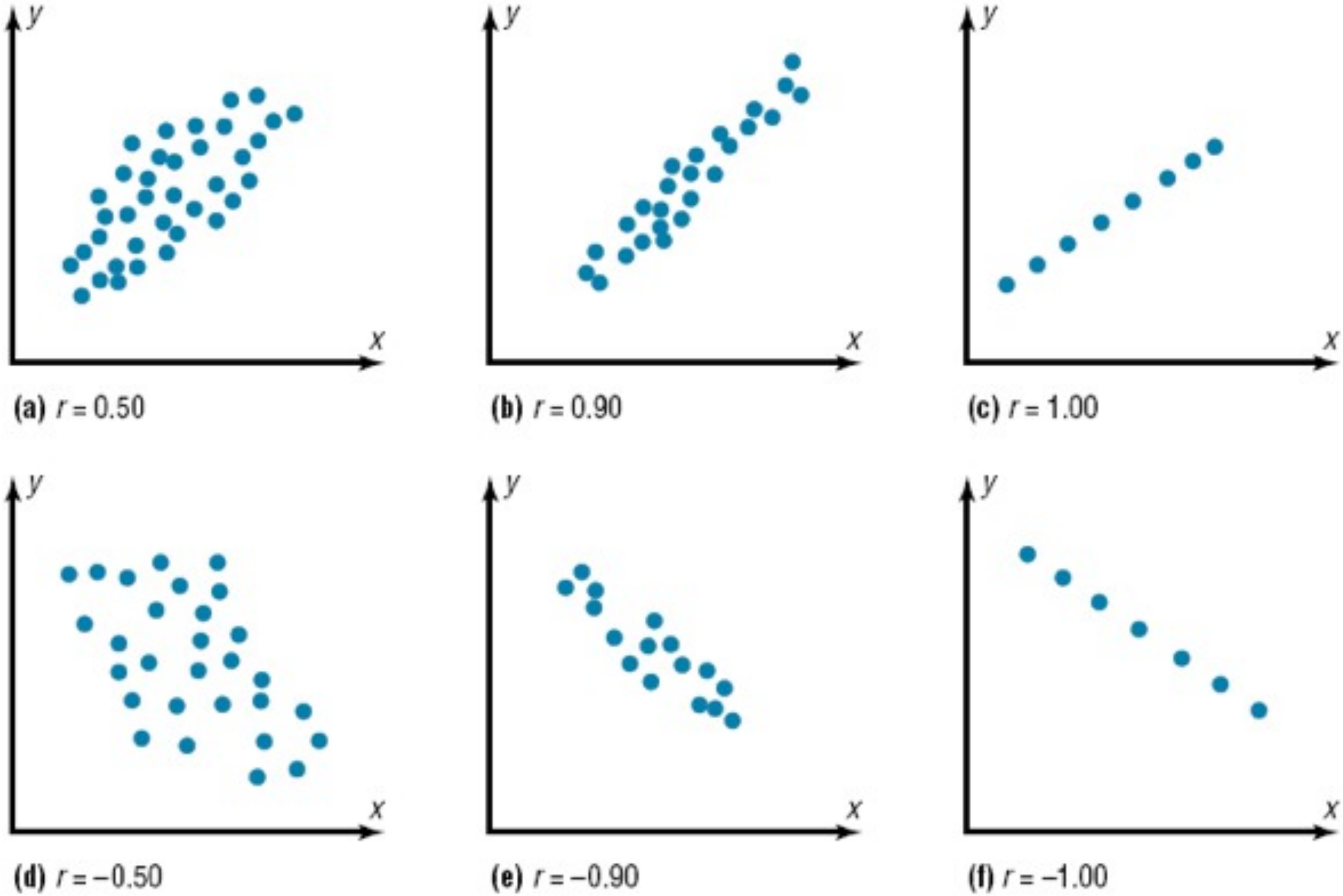
Correlation



Correlation



Correlation



Correlation Coefficient

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

Correlation Coefficient

The formula for the correlation coefficient is

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

where *n* is the number of data pairs.

Rounding Rule: Round to three decimal places.

Chapter 10
Correlation and Regression

Section 10-1

Example 10-4

Page #540

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0			
B	29.0	3.9			
C	20.8	2.1			
D	19.1	2.8			
E	13.4	1.4			
F	8.5	1.5			

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00		
B	29.0	3.9	113.10		
C	20.8	2.1	43.68		
D	19.1	2.8	53.48		
E	13.4	1.4	18.76		
F	8.5	1.5	2.75		

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	
B	29.0	3.9	113.10	841.00	
C	20.8	2.1	43.68	432.64	
D	19.1	2.8	53.48	364.81	
E	13.4	1.4	18.76	179.56	
F	8.5	1.5	2.75	72.25	

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	49.00
B	29.0	3.9	113.10	841.00	15.21
C	20.8	2.1	43.68	432.64	4.41
D	19.1	2.8	53.48	364.81	7.84
E	13.4	1.4	18.76	179.56	1.96
F	8.5	1.5	2.75	72.25	2.25

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	49.00
B	29.0	3.9	113.10	841.00	15.21
C	20.8	2.1	43.68	432.64	4.41
D	19.1	2.8	53.48	364.81	7.84
E	13.4	1.4	18.76	179.56	1.96
F	8.5	1.5	2.75	72.25	2.25
	$\Sigma x =$ 153.8				

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	49.00
B	29.0	3.9	113.10	841.00	15.21
C	20.8	2.1	43.68	432.64	4.41
D	19.1	2.8	53.48	364.81	7.84
E	13.4	1.4	18.76	179.56	1.96
F	8.5	1.5	2.75	72.25	2.25
	$\Sigma x =$ 153.8	$\Sigma y =$ 18.7			

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	49.00
B	29.0	3.9	113.10	841.00	15.21
C	20.8	2.1	43.68	432.64	4.41
D	19.1	2.8	53.48	364.81	7.84
E	13.4	1.4	18.76	179.56	1.96
F	8.5	1.5	2.75	72.25	2.25
	$\Sigma x =$ 153.8	$\Sigma y =$ 18.7	$\Sigma xy =$ 682.77		

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	49.00
B	29.0	3.9	113.10	841.00	15.21
C	20.8	2.1	43.68	432.64	4.41
D	19.1	2.8	53.48	364.81	7.84
E	13.4	1.4	18.76	179.56	1.96
F	8.5	1.5	2.75	72.25	2.25
	$\Sigma x =$ 153.8	$\Sigma y =$ 18.7	$\Sigma xy =$ 682.77	$\Sigma x^2 =$ 5859.26	

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

Company	Cars x (in 10,000s)	Income y (in billions)	xy	x^2	y^2
A	63.0	7.0	441.00	3969.00	49.00
B	29.0	3.9	113.10	841.00	15.21
C	20.8	2.1	43.68	432.64	4.41
D	19.1	2.8	53.48	364.81	7.84
E	13.4	1.4	18.76	179.56	1.96
F	8.5	1.5	2.75	72.25	2.25
	$\Sigma x =$ 153.8	$\Sigma y =$ 18.7	$\Sigma xy =$ 682.77	$\Sigma x^2 =$ 5859.26	$\Sigma y^2 =$ 80.67

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

$\Sigma x = 153.8$, $\Sigma y = 18.7$, $\Sigma xy = 682.77$, $\Sigma x^2 = 5859.26$,
$\Sigma y^2 = 80.67$, $n = 6$

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

$\Sigma x = 153.8$, $\Sigma y = 18.7$, $\Sigma xy = 682.77$, $\Sigma x^2 = 5859.26$,
$\Sigma y^2 = 80.67$, $n = 6$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

$\Sigma x = 153.8$, $\Sigma y = 18.7$, $\Sigma xy = 682.77$, $\Sigma x^2 = 5859.26$,
$\Sigma y^2 = 80.67$, $n = 6$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

$$r = \frac{(6)(682.77) - (153.8)(18.7)}{\sqrt{\left[(6)(5859.26) - (153.8)^2\right]\left[(6)(80.67) - (18.7)^2\right]}}$$

Example 10-4: Car Rental Companies

Compute the correlation coefficient for the data in Example 10–1.

$\Sigma x = 153.8$, $\Sigma y = 18.7$, $\Sigma xy = 682.77$, $\Sigma x^2 = 5859.26$,
$\Sigma y^2 = 80.67$, $n = 6$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

$$r = \frac{(6)(682.77) - (153.8)(18.7)}{\sqrt{\left[(6)(5859.26) - (153.8)^2\right]\left[(6)(80.67) - (18.7)^2\right]}}$$

$r = 0.982$ (strong positive relationship)

Chapter 10
Correlation and Regression

Section 10-1

Example 10-5

Page #541

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82			
B	2	86			
C	15	43			
D	9	74			
E	12	58			
F	5	90			
G	8	78			

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492		
B	2	86	172		
C	15	43	645		
D	9	74	666		
E	12	58	696		
F	5	90	450		
G	8	78	624		

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492	36	
B	2	86	172	4	
C	15	43	645	225	
D	9	74	666	81	
E	12	58	696	144	
F	5	90	450	25	
G	8	78	624	64	

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492	36	6,724
B	2	86	172	4	7,396
C	15	43	645	225	1,849
D	9	74	666	81	5,476
E	12	58	696	144	3,364
F	5	90	450	25	8,100
G	8	78	624	64	6,084

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, *x*	Final Grade *y* (pct.)	*xy*	x^2	y^2
A	6	82	492	36	6,724
B	2	86	172	4	7,396
C	15	43	645	225	1,849
D	9	74	666	81	5,476
E	12	58	696	144	3,364
F	5	90	450	25	8,100
G	8	78	624	64	6,084
	$\Sigma x = 57$				

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492	36	6,724
B	2	86	172	4	7,396
C	15	43	645	225	1,849
D	9	74	666	81	5,476
E	12	58	696	144	3,364
F	5	90	450	25	8,100
G	8	78	624	64	6,084
	$\Sigma x =$ 57	$\Sigma y =$ 511			

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492	36	6,724
B	2	86	172	4	7,396
C	15	43	645	225	1,849
D	9	74	666	81	5,476
E	12	58	696	144	3,364
F	5	90	450	25	8,100
G	8	78	624	64	6,084
	$\Sigma x =$ 57	$\Sigma y =$ 511	$\Sigma xy =$ 3745		

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492	36	6,724
B	2	86	172	4	7,396
C	15	43	645	225	1,849
D	9	74	666	81	5,476
E	12	58	696	144	3,364
F	5	90	450	25	8,100
G	8	78	624	64	6,084
	$\Sigma x =$ 57	$\Sigma y =$ 511	$\Sigma xy =$ 3745	$\Sigma x^2 =$ 579	

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

Student	Number of absences, x	Final Grade y (pct.)	xy	x^2	y^2
A	6	82	492	36	6,724
B	2	86	172	4	7,396
C	15	43	645	225	1,849
D	9	74	666	81	5,476
E	12	58	696	144	3,364
F	5	90	450	25	8,100
G	8	78	624	64	6,084
	$\Sigma x =$ 57	$\Sigma y =$ 511	$\Sigma xy =$ 3745	$\Sigma x^2 =$ 579	$\Sigma y^2 =$ 38,993

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

$\Sigma x = 57$, $\Sigma y = 511$, $\Sigma xy = 3745$, $\Sigma x^2 = 579$,
$\Sigma y^2 = 38{,}993$, $n = 7$

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

$\Sigma x = 57$, $\Sigma y = 511$, $\Sigma xy = 3745$, $\Sigma x^2 = 579$,
$\Sigma y^2 = 38{,}993$, $n = 7$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

$\Sigma x = 57$, $\Sigma y = 511$, $\Sigma xy = 3745$, $\Sigma x^2 = 579$,
$\Sigma y^2 = 38{,}993$, $n = 7$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

$$r = \frac{(7)(3745) - (57)(511)}{\sqrt{\left[(7)(579) - (57)^2\right]\left[(7)(38{,}993) - (511)^2\right]}}$$

Example 10-5: Absences/Final Grades

Compute the correlation coefficient for the data in Example 10–2.

$\Sigma x = 57$, $\Sigma y = 511$, $\Sigma xy = 3745$, $\Sigma x^2 = 579$,
$\Sigma y^2 = 38{,}993$, $n = 7$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

$$r = \frac{(7)(3745) - (57)(511)}{\sqrt{\left[(7)(579) - (57)^2\right]\left[(7)(38{,}993) - (511)^2\right]}}$$

$r = -0.944$ (strong negative relationship)

Chapter 10
Correlation and Regression

Section 10-1

Example 10-6

Page #542

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48			
B	0	8			
C	2	32			
D	5	64			
E	8	10			
F	5	32			
G	10	56			
H	2	72			
I	1	48			

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144		
B	0	8	0		
C	2	32	64		
D	5	64	320		
E	8	10	80		
F	5	32	160		
G	10	56	560		
H	2	72	144		
I	1	48	48		

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	
B	0	8	0	0	
C	2	32	64	4	
D	5	64	320	25	
E	8	10	80	64	
F	5	32	160	25	
G	10	56	560	100	
H	2	72	144	4	
I	1	48	48	1	

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	2,304
B	0	8	0	0	64
C	2	32	64	4	1,024
D	5	64	320	25	4,096
E	8	10	80	64	100
F	5	32	160	25	1,024
G	10	56	560	100	3,136
H	2	72	144	4	5,184
I	1	48	48	1	2,304

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	2,304
B	0	8	0	0	64
C	2	32	64	4	1,024
D	5	64	320	25	4,096
E	8	10	80	64	100
F	5	32	160	25	1,024
G	10	56	560	100	3,136
H	2	72	144	4	5,184
I	1	48	48	1	2,304
	$\Sigma x = 36$				

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	2,304
B	0	8	0	0	64
C	2	32	64	4	1,024
D	5	64	320	25	4,096
E	8	10	80	64	100
F	5	32	160	25	1,024
G	10	56	560	100	3,136
H	2	72	144	4	5,184
I	1	48	48	1	2,304
	$\Sigma x =$ 36	$\Sigma y =$ 370			

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	2,304
B	0	8	0	0	64
C	2	32	64	4	1,024
D	5	64	320	25	4,096
E	8	10	80	64	100
F	5	32	160	25	1,024
G	10	56	560	100	3,136
H	2	72	144	4	5,184
I	1	48	48	1	2,304
	$\Sigma x =$ 36	$\Sigma y =$ 370	$\Sigma xy =$ 1,520		

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	2,304
B	0	8	0	0	64
C	2	32	64	4	1,024
D	5	64	320	25	4,096
E	8	10	80	64	100
F	5	32	160	25	1,024
G	10	56	560	100	3,136
H	2	72	144	4	5,184
I	1	48	48	1	2,304
	$\Sigma x =$ 36	$\Sigma y =$ 370	$\Sigma xy =$ 1,520	$\Sigma x^2 =$ 232	

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

Subject	Hours, x	Amount y	xy	x^2	y^2
A	3	48	144	9	2,304
B	0	8	0	0	64
C	2	32	64	4	1,024
D	5	64	320	25	4,096
E	8	10	80	64	100
F	5	32	160	25	1,024
G	10	56	560	100	3,136
H	2	72	144	4	5,184
I	1	48	48	1	2,304
	$\Sigma x =$ 36	$\Sigma y =$ 370	$\Sigma xy =$ 1,520	$\Sigma x^2 =$ 232	$\Sigma y^2 =$ 19,236

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

$\Sigma x = 36$, $\Sigma y = 370$, $\Sigma xy = 1520$, $\Sigma x^2 = 232$,
$\Sigma y^2 = 19{,}236$, $n = 9$

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

$\Sigma x = 36,\ \Sigma y = 370,\ \Sigma xy = 1520,\ \Sigma x^2 = 232,$
$\Sigma y^2 = 19{,}236,\ n = 9$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

$\Sigma x = 36$, $\Sigma y = 370$, $\Sigma xy = 1520$, $\Sigma x^2 = 232$,
$\Sigma y^2 = 19{,}236$, $n = 9$

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$$

$$r = \frac{(7)(1520) - (36)(370)}{\sqrt{\left[(7)(232) - (36)^2\right]\left[(7)(19{,}236) - (370)^2\right]}}$$

Example 10-6: Exercise/Milk Intake

Compute the correlation coefficient for the data in Example 10–3.

$\Sigma x = 36$, $\Sigma y = 370$, $\Sigma xy = 1520$, $\Sigma x^2 = 232$,
$\Sigma y^2 = 19{,}236$, $n = 9$

$$r = \frac{n\left(\sum xy\right)-\left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right)-\left(\sum x\right)^2\right]\left[n\left(\sum y^2\right)-\left(\sum y\right)^2\right]}}$$

$$r = \frac{(7)(1520)-(36)(370)}{\sqrt{\left[(7)(232)-(36)^2\right]\left[(7)(19,236)-(370)^2\right]}}$$

$r = 0.067$ (very weak relationship)

Hypothesis Testing

- In hypothesis testing, one of the following is true:

Hypothesis Testing

- In hypothesis testing, one of the following is true:

  H_0: $\rho = 0$ This null hypothesis means that there is no correlation between the *x* and *y* variables in the population.

Hypothesis Testing

- In hypothesis testing, one of the following is true:

  H_0: $\rho = 0$ This null hypothesis means that there is no correlation between the *x* and *y* variables in the population.

Hypothesis Testing

- In hypothesis testing, one of the following is true:

  H_0: $\rho = 0$ This null hypothesis means that there is no correlation between the *x* and *y* variables in the population.

  H_1: $\rho \neq 0$ This alternative hypothesis means that there is a significant correlation between the variables in the population.

t Test for the Correlation Coefficient

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

with degrees of freedom equal to $n - 2$.

Chapter 10
Correlation and Regression

Section 10-1

Example 10-7

Page #544

Example 10-7: Car Rental Companies

Test the significance of the correlation coefficient found in Example 10–4. Use $\alpha = 0.05$ and $r = 0.982$.

Example 10-7: Car Rental Companies

Test the significance of the correlation coefficient found in Example 10–4. Use $\alpha = 0.05$ and $r = 0.982$.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

Example 10-7: Car Rental Companies

Test the significance of the correlation coefficient found in Example 10–4. Use $\alpha = 0.05$ and $r = 0.982$.

Step 1: State the hypotheses.

$H_0: \rho = 0$ and $H_1: \rho \neq 0$

Example 10-7: Car Rental Companies

Test the significance of the correlation coefficient found in Example 10–4. Use $\alpha = 0.05$ and $r = 0.982$.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

Step 2: Find the critical value.

Example 10-7: Car Rental Companies

Test the significance of the correlation coefficient found in Example 10–4. Use $\alpha = 0.05$ and $r = 0.982$.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

Step 2: Find the critical value.

Since $\alpha = 0.05$ and there are $6 - 2 = 4$ degrees of freedom, the critical values obtained from Table F are ±2.776.

Example 10-7: Car Rental Companies

Example 10-7: Car Rental Companies

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

Example 10-7: Car Rental Companies

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-\left(0.982\right)^2}} = 10.4$$

Example 10-7: Car Rental Companies

Step 3: Compute the test value.

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-(0.982)^2}} = 10.4$$

Example 10-7: Car Rental Companies

Step 3: Compute the test value.

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-\left(0.982\right)^2}} = 10.4$$

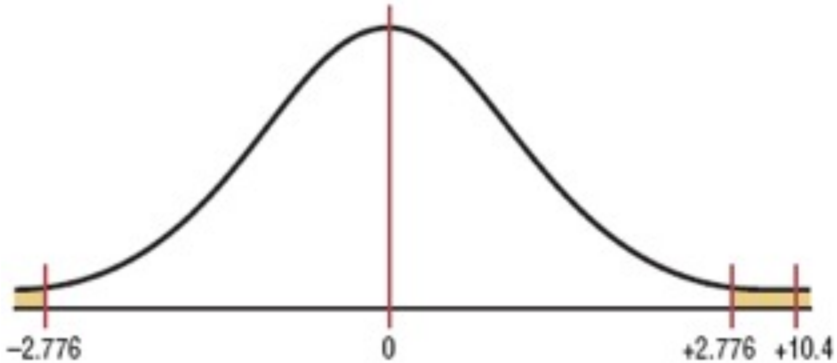
Step 4: Make the decision.

Example 10-7: Car Rental Companies

Step 3: Compute the test value.

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-(0.982)^2}} = 10.4$$

Step 4: Make the decision.



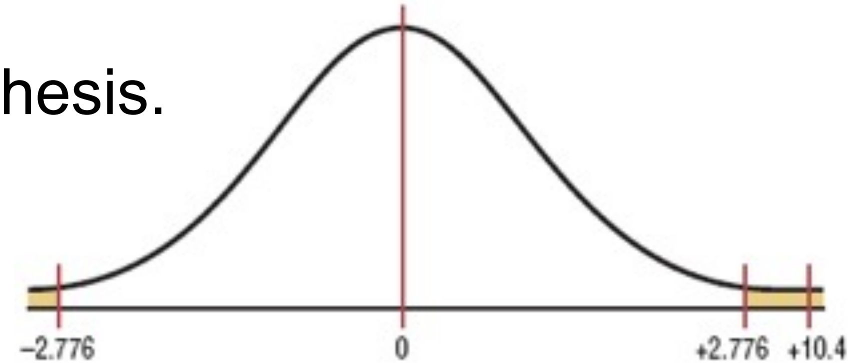
Example 10-7: Car Rental Companies

Step 3: Compute the test value.

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-(0.982)^2}} = 10.4$$

Step 4: Make the decision.

Reject the null hypothesis.



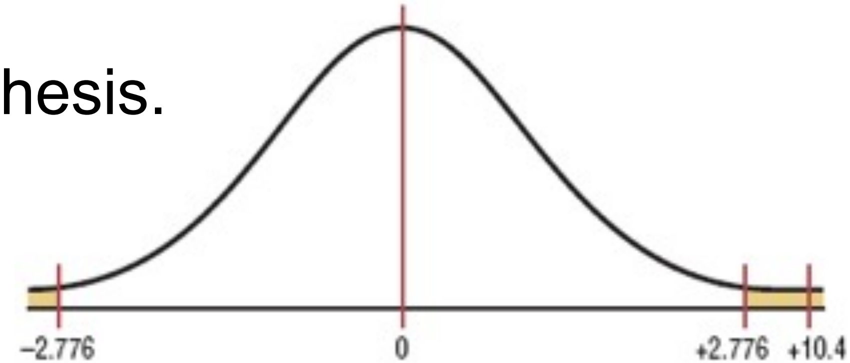
Example 10-7: Car Rental Companies

Step 3: Compute the test value.

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.982\sqrt{\frac{6-2}{1-(0.982)^2}} = 10.4$$

Step 4: Make the decision.

Reject the null hypothesis.



Step 5: Summarize the results.

There is a significant relationship between the number of cars a rental agency owns and its annual income.

Chapter 10
Correlation and Regression

Section 10-1

Example 10-8

Page #545

Example 10-8: Car Rental Companies

Using Table I, test the significance of the correlation coefficient r = 0.067, from Example 10–6, at α = 0.01.

Example 10-8: Car Rental Companies

Using Table I, test the significance of the correlation coefficient $r = 0.067$, from Example 10–6, at $\alpha = 0.01$.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

Example 10-8: Car Rental Companies

Using Table I, test the significance of the correlation coefficient $r = 0.067$, from Example 10–6, at $\alpha = 0.01$.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

There are $9 - 2 = 7$ degrees of freedom. The value in Table I when $\alpha = 0.01$ is 0.798.

Example 10-8: Car Rental Companies

Using Table I, test the significance of the correlation coefficient $r = 0.067$, from Example 10–6, at $\alpha = 0.01$.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

There are 9 – 2 = 7 degrees of freedom. The value in Table I when $\alpha = 0.01$ is 0.798.

For a significant relationship, r must be greater than 0.798 or less than -0.798. Since $r = 0.067$, do not reject the null.

Example 10-8: Car Rental Companies

Using Table I, test the significance of the correlation coefficient r = 0.067, from Example 10–6, at α = 0.01.

Step 1: State the hypotheses.

H_0: $\rho = 0$ and H_1: $\rho \neq 0$

There are 9 – 2 = 7 degrees of freedom. The value in Table I when α = 0.01 is 0.798.

For a significant relationship, r must be greater than 0.798 or less than -0.798. Since r = 0.067, do not reject the null.

Hence, there is not enough evidence to say that there is a significant linear relationship between the variables.

Example 10-8: Car Rental Companies

Using Table I, test the significance of the correlation coefficient r = 0.067, from Example 10–6, at α = 0.01.

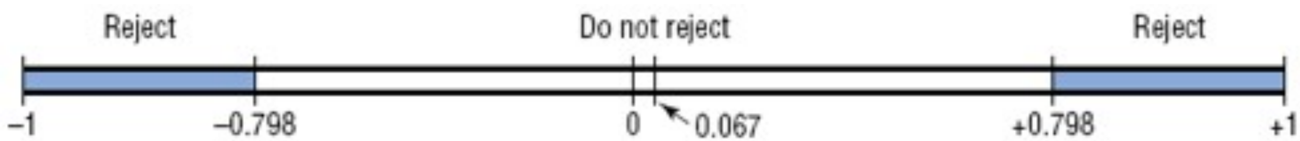
Step 1: State the hypotheses.

H_0: ρ = 0 and H_1: $\rho \neq 0$

There are 9 – 2 = 7 degrees of freedom. The value in Table I when α = 0.01 is 0.798.

For a significant relationship, r must be greater than 0.798 or less than -0.798. Since r = 0.067, do not reject the null.

Hence, there is not enough evidence to say that there is a significant linear relationship between the variables.



Possible Relationships Between Variables

When the null hypothesis has been rejected for a specific a value, any of the following five possibilities can exist.

Possible Relationships Between Variables

When the null hypothesis has been rejected for a specific a value, any of the following five possibilities can exist.

1. There is a *direct cause-and-effect* relationship between the variables. That is, *x* causes *y*.

Possible Relationships Between Variables

When the null hypothesis has been rejected for a specific a value, any of the following five possibilities can exist.

1. There is a *direct cause-and-effect* relationship between the variables. That is, *x* causes *y*.
2. There is a *reverse cause-and-effect* relationship between the variables. That is, *y* causes *x*.

Possible Relationships Between Variables

When the null hypothesis has been rejected for a specific a value, any of the following five possibilities can exist.

1. There is a *direct cause-and-effect* relationship between the variables. That is, *x* causes *y*.
2. There is a *reverse cause-and-effect* relationship between the variables. That is, *y* causes *x*.
3. The relationship between the variables may be *caused by a third variable*.

Possible Relationships Between Variables

When the null hypothesis has been rejected for a specific a value, any of the following five possibilities can exist.

1. There is a *direct cause-and-effect* relationship between the variables. That is, x causes y.
2. There is a *reverse cause-and-effect* relationship between the variables. That is, y causes x.
3. The relationship between the variables may be *caused by a third variable*.
4. There may be a *complexity of interrelationships* among many variables.

Possible Relationships Between Variables

When the null hypothesis has been rejected for a specific a value, any of the following five possibilities can exist.

1. There is a *direct cause-and-effect* relationship between the variables. That is, *x* causes *y*.
2. There is a *reverse cause-and-effect* relationship between the variables. That is, *y* causes *x*.
3. The relationship between the variables may be *caused by a third variable*.
4. There may be a *complexity of interrelationships* among many variables.
5. The relationship may be *coincidental*.

Possible Relationships Between Variables

1. There is a *reverse cause-and-effect* relationship between the variables. That is, *y* causes *x.*

Possible Relationships Between Variables

1. There is a *reverse cause-and-effect* relationship between the variables. That is, *y* causes *x.*

   For example,

   - water causes plants to grow
   - poison causes death
   - heat causes ice to melt

Possible Relationships Between Variables

Possible Relationships Between Variables

2. There is a *reverse cause-and-effect* relationship between the variables. That is, *y* causes *x.*

For example,

- Suppose a researcher believes excessive coffee consumption causes nervousness, but the researcher fails to consider that the reverse situation may occur. That is, it may be that an extremely nervous person craves coffee to calm his or her nerves.

Possible Relationships Between Variables

Possible Relationships Between Variables

3. The relationship between the variables may be *caused by a third variable*.

For example,

- If a statistician correlated the number of deaths due to drowning and the number of cans of soft drink consumed daily during the summer, he or she would probably find a significant relationship. However, the soft drink is not necessarily responsible for the deaths, since both variables may be related to heat and humidity.

Possible Relationships Between Variables

Possible Relationships Between Variables

4. There may be a *complexity of interrelationships* among many variables.

   For example,

   - A researcher may find a significant relationship between students' high school grades and college grades. But there probably are many other variables involved, such as IQ, hours of study, influence of parents, motivation, age, and instructors.

Possible Relationships Between Variables

Possible Relationships Between Variables

5. The relationship may be *coincidental.*

   For example,

   - A researcher may be able to find a significant relationship between the increase in the number of people who are exercising and the increase in the number of people who are committing crimes. But common sense dictates that any relationship between these two values must be due to coincidence.