Exercises

2.4.1 To determine if there is a relationship between shoe length and height of a person, the author had ten of his students measure their shoe length (to the nearest quarter of an inch) and height (to the nearest half inch). The resulting data are show in Table 2.5. Graph Height vs Shoe Length and fit a straight line. How well does this model fit the data?

Shoe Length	9	10	10.5	11	11.5	11.75	12	12.5	12.75	13
Height	62	64	64.5	69	70	73	72	75	74	77

Table 2.5

2.4.2 Table 2.6 shows the price (in dollars) of several name-brand items and their generic equivalents at a large retail store (data collected by Amanda Schroeder, 2010).

Name	7.68	9.56	7.62	1.58	2.98	3.58	6.54	8.97	4.06	6.46
Generic	4.47	2.97	3.34	0.67	3.12	2.98	5.34	6.97	1.76	4.12

Table 2.6

a. For each product, calculate the percent savings if the generic product is purchased instead of the name-brand.

b. Fit a straight line to the graph of percent savings vs price of name-brand product. How well does this model fit the data?

2.4.3 Suppose a biologist records the number of pulses per second of the chirps of a cricket at different temperatures (in °F). The data collected is shown in Table 2.7.

Temperature	72	73	89	75	93	85	79	97	86	91
Pulses/sec	16	16.2	21.2	16.5	20	18	16.75	19.25	18.25	18.5

Table 2.7

a. Fit a straight line to this data (where temperature is on the x-axis). How well does the model fit the data?

b. What is the slope of this line? What does the sign of the slope tell you about the relationship between pulses/sec and temperature?

2.4.4 Table 2.8 gives the number of manatee deaths in Florida believed to be caused by watercraft for the years 1983 - 2005 where Year 1 corresponds to 1983 (data from the Florida Fish and Wildlife Conservation Commission, Marine Mammal Pathobiology Laboratory). Fit a straight line to this data (where year is on the x-axis). How well does the model fit the data? Do you notice any characteristics of the data that make you question whether a linear (straight line) model is the "best" type of model?

Year	1	2	3	4	5	6	7	8	9	10	11	12
Deaths	15	24	44	44	39	43	50	47	53	38	35	49
Year	13	14	15	16	17	18	19	20	21	22	23	24
Deaths	42	60	54	67	82	78	81	95	73	69	79	92

Table 2.8

2.4.5 Starting with equation (2.6), show that a formula for b in terms of m is $b = \bar{y} - m\bar{x}$ where

$$\bar{y} = \frac{1}{n}\sum y_i \quad \text{and} \quad \bar{x} = \frac{1}{n}\sum x_i$$

are the averages of the y- and x-values, respectively.

2.4.6 Starting with equations (2.5) and (2.6), derive the formula for m in equation (2.7). (**Hint**: Multiply equation (2.5) by $-n$ and (2.6) by $\sum x_i$. Add the resulting equations and solve this result for m.)

2.4.7 As we saw in Section 2.2, often times we want to fit a straight line that goes through the origin to a set of data. This means we want the y-intercept of the straight line to be 0. So the equation of the line is simply $y = mx$. In this case, the least–squares criterion says that we want to minimize

$$S = \sum_{i=1}^{n} (y_i - mx_i)^2.$$

a. Take the derivative of this equation with respect to m, set it equal to 0, and solve it for m to find a formula for m in terms of the x- and y-coordinates.

b. Implement this formula in Excel and use it to fit a straight line through the origin to the data in Table 2.1 on page 15. How does the slope of this line compare to the slope of the line found in Example 2.2.1?

c. This formula can be implemented by adding a trendline to the data and selecting **Set intercept =** under **Options**. Do this to the data and compare the slope of the trendline to the slope calculated by your formula. Are they indeed equal?

2.4.8 Generalize your result in 2.4.7. Suppose you have specified a value of b, say $b = b_0$, so that the model is now $y = mx + b_0$. In this case, the least-squares criterion says that we want to minimize

$$S = \sum_{i=1}^{n} [y_i - (mx_i + b_0)]^2$$

a. Take the derivative of this equation with respect to m, set it equal to 0, and solve it for m to find a formula for m in terms of the x- and y-coordinates and b_0.

b. Suppose you hang a bucket from a spring that stretches it 15 mm from its natural length. You then fill the bucket with different weights of sand and record the total distance stretched as recorded below.

Weight (N)	10	15	20	25	30	35
Distance (mm)	57.4	67.6	71.6	104.5	106.9	136.0

Because the spring was stretched before the first amount of sand was added, Hooke's law predicts that the relationship between the distance stretched, D, and the weight of sand, W, is

$$D = mW + 15.$$

Create a graph of the data to test if this prediction is reasonable. If it is, use your formula from part a. to estimate the value of m.

c. We could also estimate the value of m by transforming the data by subtracting 15 from each distance measurement, and then fitting a straight line through the origin to the transformed data. The slope of this line is m. Do this and compare the value of m to that found in part b.

2.4.9 Consider the problem of fitting a model of the form $y = a$ to a set of data. Show that when using the least-squares criterion, the appropriate formula for the parameter a is $a = \frac{1}{n}\sum_{i-1}^{n} y_i$.

2.4.10 When using Chebyshev's criterion to fit a straight line to data, we find the values of m and b that minimize the number

$$C = \text{Maximum of } \{|y_i - (mx_i + b)| \,:\, i = 1, 2, \ldots, n\}.$$

Consider the problem of fitting a straight line to the data in Figure 2.14, but set $m = 0.694$. Use a scroll bar to estimate the value of b that minimizes the value of C in Chebyshev's criterion. Graph the resulting line on the data and compare it to the line found using the least-squares criterion. Does Chebyshev's criterion give the exact same value of b as the least-squares criterion?

2.4.11 When using the sum of the distances criterion to fit a straight line to data, we find the values of m and b that minimize the number

$$A = \sum_{i=1}^{n} |y_i - (mx_i + b)|.$$

Consider the problem of fitting a straight line to the data in Figure 2.14, but set $b = 1.69$. Use a scroll bar to estimate the value of m that minimizes the value of A in this criterion. Graph the resulting line on the data and compare it to the line found using the least-squares criterion. Does this criterion give the exact same value of m as the least-squares criterion?

2.4.12 Sports Manufacturing Inc. manufactures footballs, basketballs, and soccer balls. Each week the company manufactures a different type of ball in varying quantities. Manufacturing costs fall into two different categories: start-up and unit. Start-up costs are costs necessary to begin production

of a particular product (retool machinery, etc.). Unit costs ($/unit) are the costs associated with manufacturing individual units (labor, materials, etc.).

The table below shows data for 15 weeks of production. The column "Total Cost" gives the total cost to produce the given number of units of that type of ball in a week. Your goal is to model the total cost, estimate the start-up and unit costs for each product, and implement the model.

Footballs		**Basketballs**		**Soccer balls**	
Units	**Total Cost**	**Units**	**Total Cost**	**Units**	**Total Cost**
2222	3125	962	1520	2481	4300
2263	3250	2246	2850	1825	3190
1267	1955	2430	2990	2238	3930
2177	3120	1395	1920	949	1890
2266	3090	2405	2750	1250	2350

a. Define variables and create a model for the total cost for producing a given number of units of each product in terms of the start-up and unit costs. List the assumptions you make.

b. Estimate the start-up and unit costs for each product.

c. Create a spreadsheet in which a user can easily input production data, such as that shown above, along with the number of units of a product that are planned for production in a given week and see an estimated total production cost for that week. Make sure the spreadsheet is logical and easy to use.

2.5 Geometric Similarity

Shapes such as circles and rectangles are easy to work with. We can calculate the area and volume of objects with these shapes using very simple formulas. Real world objects rarely come in these simple forms. This necessitates some simplifying assumptions. Geometric similarity is one such assumption.

Definition 2.5.1 Two objects are **geometrically similar** if the following two conditions are met:

1. There is a one-to-one correspondence between points of the objects (i.e. the two objects have the same "shape").
2. The ratio of distances between corresponding points is the same for all pairs of points.

In simpler terms, two objects are geometrically similar if one is a scaled up or down version of the other.