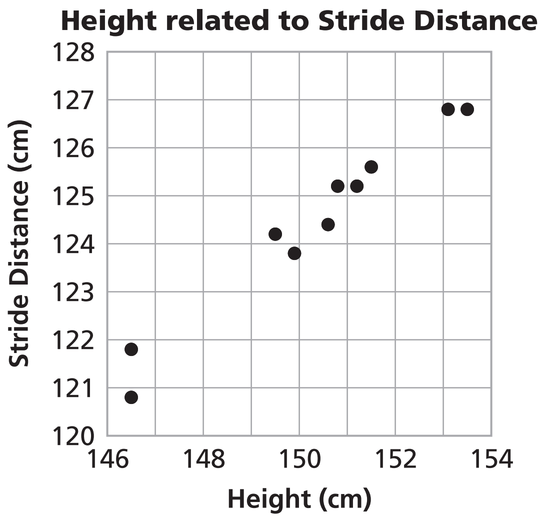
Applications

1. a. Median height is 150.7 cm. Order the 10 heights from shortest to tallest. Since 10 is even, average the two middle numbers, 150.6 cm and 150.8 cm.

   b. Median stride distance is 124.8 cm. Order the 10 stride distances from shortest to longest. Since 10 is even, average the two middle numbers, 124.4 cm and 125.2 cm.

   c. The median height is about 1.2 (1.20753205) times the median stride distance. (**Note:** For Exercise 2, you might explore with the students what a line drawn at stride distance = height ÷ 1.2 means.)

2. a. The graph indicates that, in general, taller people have longer stride distances. Knowing a person's stride will not definitively tell you that person's height, but 1.2 (stride distance) would be a good estimate. (See Exercise 1.)



   b. Identify the shortest height and then look at the corresponding stride distance; shorter people do have shorter strides.

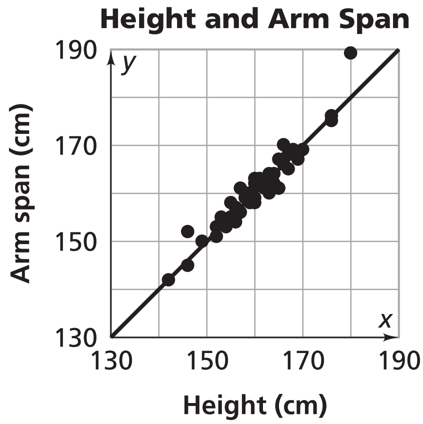
   c. 1.2(stride distance) = height

      i. 180 cm; 1.2(150) = 180

      ii. 108 cm; 1.2(90) = 108

      iii. 132 cm; 1.2(110) = 132

3. a.



   b. The line suggests that arm span is always equal to height. (**Note:** This exercise is used in Exercise 5.)

   c. The line will have equation $a = h$.

   d. i. arm span equals height.

      ii. arm span is greater than height

      iii. arm span is less than height

4. **a.** (See Figure 1.)

b. The data suggest that for jet planes the body length is consistently longer than the wingspan. For propeller planes the opposite is true.

c. If you ignore the differences between jet and propeller planes, the trend line has equation $W = 0.8L + 9.2$ and the prediction would be (40, 41.2). (**Note:** Student equations are likely to be different than the best fit equations provided in this answer.) If you separate jets from propeller planes and draw two trend lines, the predictions would be:

Jet: Trend line: $W = L - 6$; Prediction: (40, 34)

Propeller: Trend line: $W = 0.86L + 9.1$; Prediction: (40, 43.5)

d. If you ignore the differences between jet and propeller planes, the trend line has equation $W = 0.8L + 9.2$ and the prediction would be (63.5, 60). (**Note:** Student equations are likely to be different than the best fit equations provided in this answer.) If you separate jets from propeller planes and draw two trend lines, the predictions would be:

Jet: Trend line: $W = L - 6$; Prediction: (66, 60)

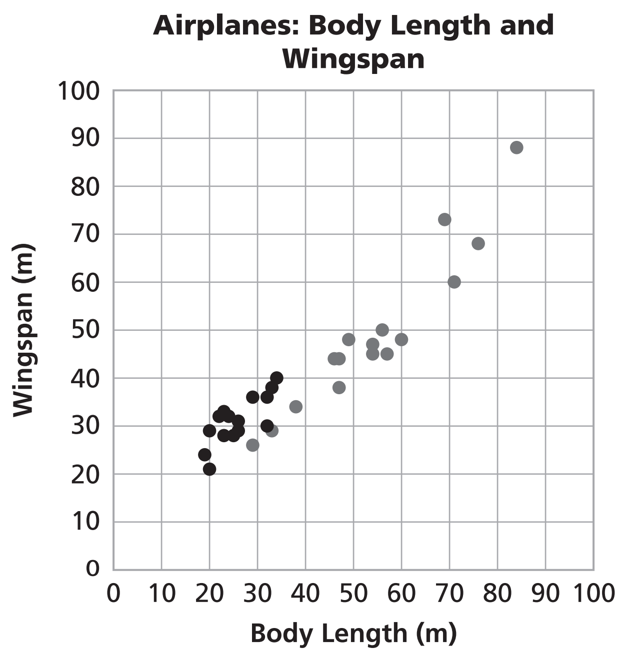
Propeller: Trend line: $W = 0.86L + 9.1$; Prediction: (59.2, 60)

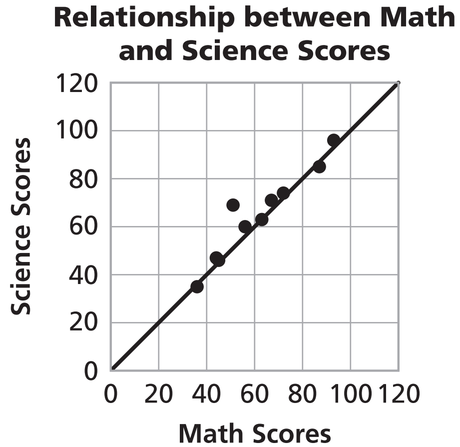
5. **a.** The equation $w = \ell$ using w for wingspan and ℓ for body length is not a good fit of the data.

b. Estimates of a linear model that is a good fit will vary. $w = 2\ell$ is a pretty good estimate. This line has y-intercept (0, 0) and slope 2, meaning that wingspan is twice body length.

c. Using the linear model from part (b), the predicted wingspan would be 120 inches for a body length of 60 inches.

Figure 3



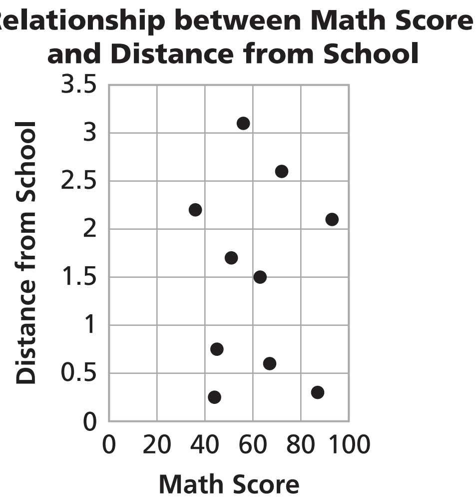
6. a. 

b. The math and science scores are similar for each student.

c. See the line drawn on the graph. (**Note:** The line $s = m$ is a good fit for the data.)

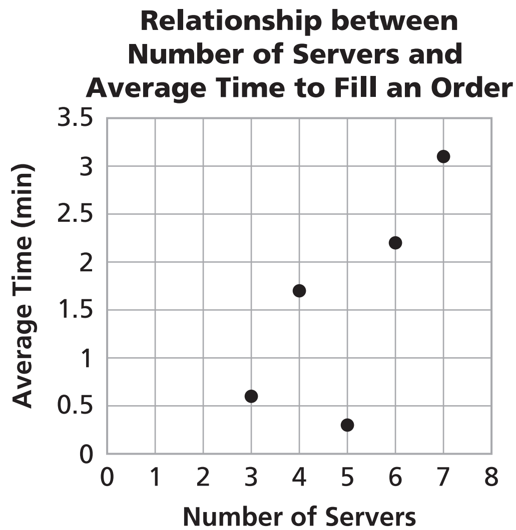
d. Student 2 appears to be an outlier, having a higher science score than would be predicted by the math score.

e. The correlation coefficient is closest to $r = 1$. (**Note:** The actual value is $r = 0.96$ but students can't estimate that value.)

7. a. 

b. There does not seem to be a relationship between math score and distance a student lives from school. The points in the scatter plot do not cluster in a linear pattern.

c. The correlation coefficient is closest to $r = 0$. (**Note:** The actual value is $r = 0.002$ but students can't estimate that value.)

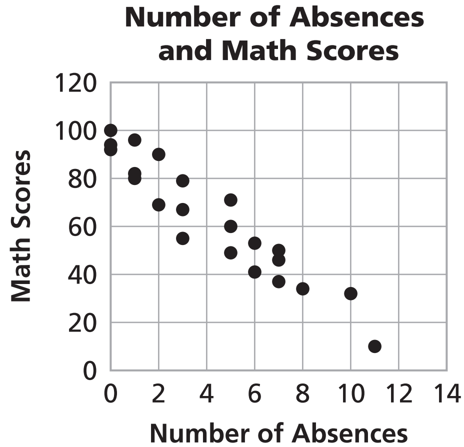
8. a. 

b. If anything, the data show a trend for higher average time when there are more servers. (**Note:** This is a very small data set.)

c. The point (5, 0.3) is an outlier. This might be because the serving time at the restaurant seems much shorter than what one expects from the trend in the data points.

d. The correlation coefficient is closest to $r = 0.5$ or 1. (**Note:** The actual value is $r = 0.76$ but students can't estimate that value.)

9. a.



b. There is a clear trend relating absences to math scores, with more absences generally associated with lower math scores.

c. Students should graph a line close to the best fit line. The line of best fit has equation $m = -7.1a + 93$, meaning that each additional absence is associated with a decrease of about 7 test score points.

d. The correlation coefficient is closest to $r = -1$. (**Note:** The actual value is $r = -0.95$ but students can't estimate that value. This indicates a strong negative association between the variables.)

10. The graph of heights for students in Class 2 has two clusters. The larger cluster is found between 136 and 140 cm. The smaller cluster is found between 144 and 148 cm. The mean of 143.8 and median of 145 are close to each other at the center of the distribution. The graph shows that a student in Class 2 has an unusual height of 163 cm. The standard deviation of 8.6 suggests that a large fraction of the heights are between 135 and 152 cm.

11. a. (See Figure 2.)

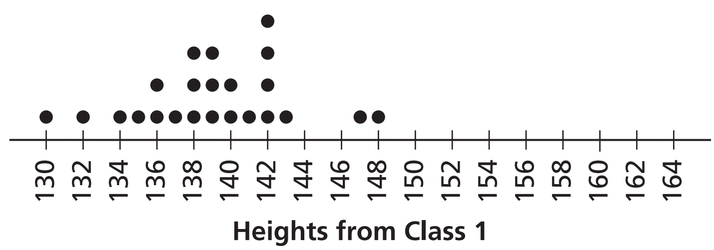
b. mean = 139.0; median = 139; range 148 − 130 = 18; standard deviation = 4.26

c. The graph of heights for students in Class 1 are clustered between 136 and 142 cm, with 17 of the 23 heights in this interval. The mean of 139 cm is the same as the median at the center of the distribution. There are four unusual heights, 130, 132, 147, and 148 cm. The standard deviation of 4.26 suggests that a large fraction of the heights will be between about 135 and 143 cm.

d. This sample of student heights in Class 2 has greater variability than the sample in Class 1.

e. Neither class data set is good for predicting the height of a typical student. The two samples have quite different mean and median heights and the data from Class 2 is very spread out which makes a typical student hard to identify.

Figure 2



12. a. Set A has mean = 10; Set B has mean = 10; Set C has mean = 10.

b. Set A has standard deviation = 2.236;
Set B has standard deviation = 0;
Set C has standard deviation = 9.

c. Set B has no standard deviation because the values are all equal to the mean. Set C has two values that are the same and smaller than any values in Set A. Set C also has two values that are the same and larger than any values in Set A. This means that Set C has the greatest standard deviation because it has the greatest spread.

13. a. Mean = \$3,955

b. Standard deviation = \$289.24

Connections

14. a. A ratio greater than 1 means arm span is greater than height. On a plot of (*h*, *s*) data and the line $s = h$, these points would be above the model line.

b. A ratio equal to 1 means height is equal to arm span. On a plot of (*h*, *s*) data and the line $s = h$, these points would be on the model line.

c. A ratio less than 1 means arm span is less than height. On a plot of (*h*, *s*) data and the line $s = h$, these points would be above the model line.

15. C; In comparing shoes and jump height, look at the clusters to describe differences of the two distributions since measures of center lead to the same conclusion. (**Note:** The case for either choice B or choice C could be made based on the mean or clustering.)

16. a. The distribution is skewed to the right when the mean is greater than the median.

b. The distribution is skewed to the left when the mean is less than the median.

c. When the mean and median are equal, the distribution is symmetrical.

17. H; The mode is 100, which is higher than either the mean or median.

18. B; $5(6.7) - 4(7.2) = 4.7$

19. H; $\frac{3(90) + 86}{4} = 89$

20. a. one possible answer: 1, 2, 3, 4

b. one possible answer: 1, 2, 3

c. one possible answer: 2, 2, 2

Extensions

21. a. $F = 0.25s + 40$

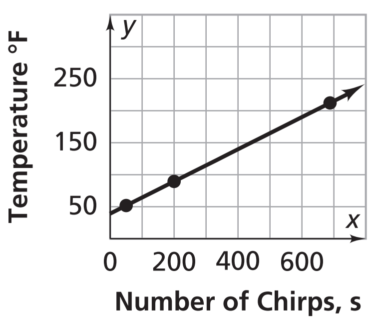
b. See graph. 0°; None (this part of the graph would not be used in this context)

50°F: about 40

100°F: about 240

212°F: about 688

Temperature versus Chirps



c. **i.** 60 chirps

ii. 120 chirps

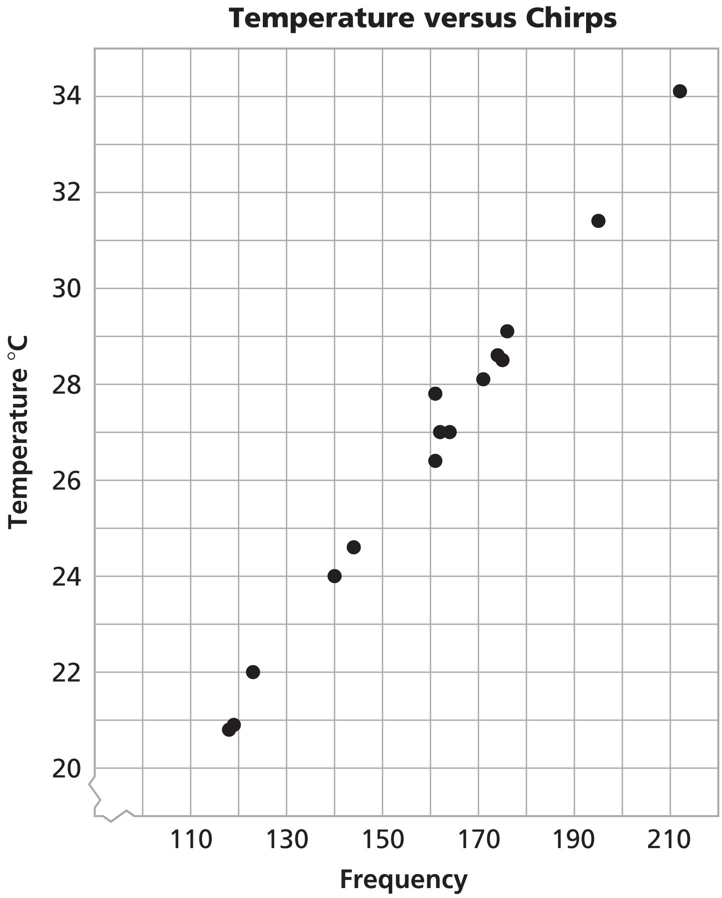
iii. 180 chirps

iv. 240 chirps

22. **a.** (See Figure 3.)

b. $T_c = \frac{5}{36}x + \frac{40}{9}$

Figure 3



c. The line is a good fit and a good model because the points cluster close to it and there are no outliers. (See Figure 4.)

23. a. The mean and the median are about the same, suggesting a roughly symmetric distribution of estimated weights.

b. The mean is less than the median, suggesting a distribution that is somewhat skewed left.

c. The correlation coefficient is closest to $r = 1$. (**Note:** The actual value is $r = 0.81$ but students can't estimate that value.)

24. a. The actual counts vary from 309 to 607 with a median of 458.5 (halfway between the 17th and 18th counts).

b. The estimates vary from 200 to 2,000 with median of 642.5 (halfway between the 17th and 18th estimates).

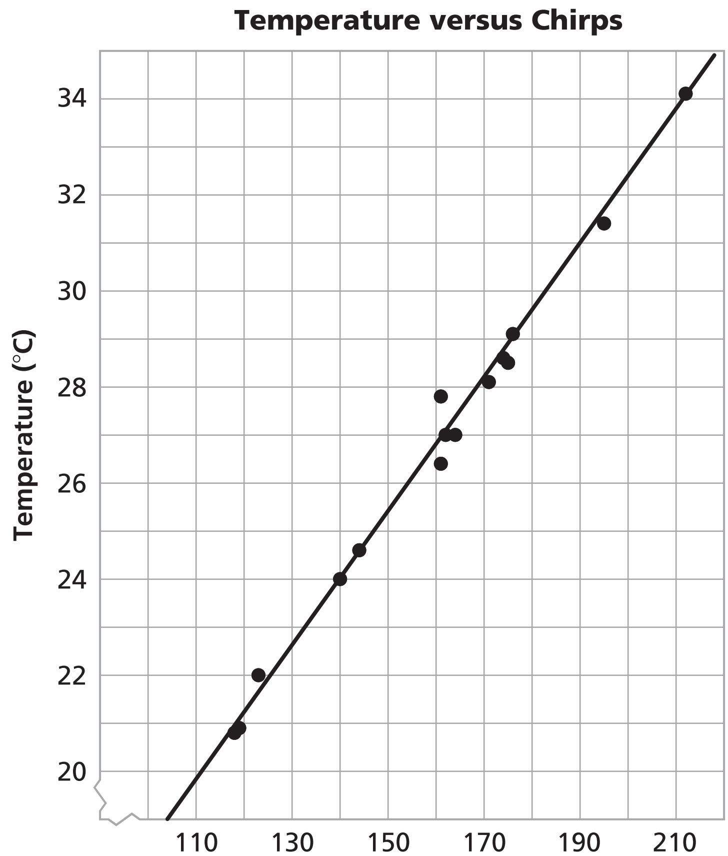
c. Points near the line represent cases where the estimates and actual counts are approximately equal.

d. Points above the line represent cases where the estimate is greater than the actual count.

e. Points below the line represent cases where the estimate is less than the actual count.

f. Answers will vary. Most students made poor estimates since there are few points on or near the estimate = actual count line. The range of estimates is much greater than the range of actual counts. The median of estimates is much greater than the median of actual counts.

Figure 4



g. A correlation coefficient for the variables of estimates and the actual counts is approximately zero ($r = -0.13$).

h. Answers will vary. Change the *x*-axis. Make the maximum 800 or 1,000 and change the increment from 100 to 50. This spreads out the points horizontally. If the *y*-axis is lengthened, then increments could be changed from 100 to 50 and this would spread the points out as well.

25. C; As the number of MP3s downloaded increases, the amount of unused space decreases.

26. a. The graph is clearly an upward curve. The rates of change in weight are, for points between 50 and 75, about 3 pounds per inch and, for points between 125 and 150, about 12 pounds per inch. (**Note:** Using the slope formula $m = \frac{y_2 - y_1}{x_2 - x_1}$, then $\frac{125 - 50}{75 - 50} =$ 3 pounds per inch. $\frac{750 - 450}{150 - 125} =$ 12 pounds per inch.)

b. Students may choose a quadratic or a cubic family of functions. These graphs have roughly the same shape as the graph of the model of the relationship between pumpkin circumference and pumpkin weight. Testing values for *k*, might lead to a simple quadratic like $w = 0.03c^2$. From a geometric perspective, weight is proportional to the cube of circumference making it directly proportional to volume or the cube of three linear dimensions. Considering that the core of a pumpkin is much less dense than the outer shell, the quadratic relationship is supported by the proportional relationship between surface area of the pumpkin (or the square of the circumference) and the pumpkin weight.