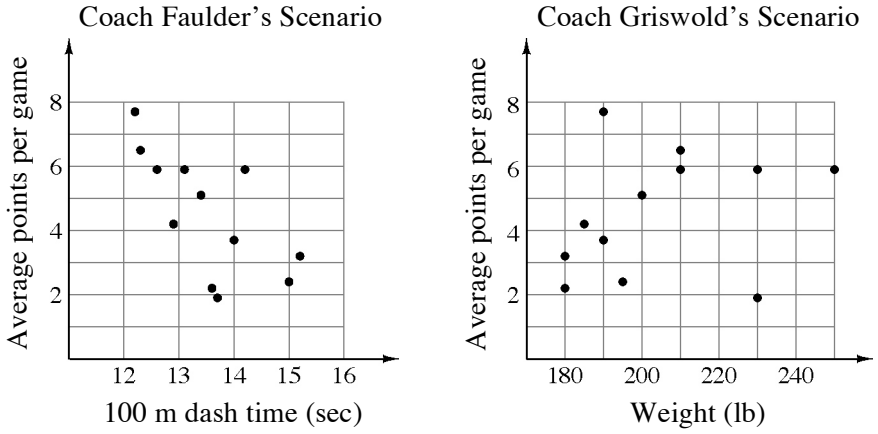
Statistics: Chapter 2 Solutions

Lesson 2.1.1

2-1. The 100 meter dash time is a better predictor of points scored, seemingly; the fastest runners seem to score a lot of points. This can be hard to see without a scatterplot.

2-2. Griswold: explanatory is weight, response is points scored. Faulder: explanatory is 100 m time, response is points scored.

2-3.

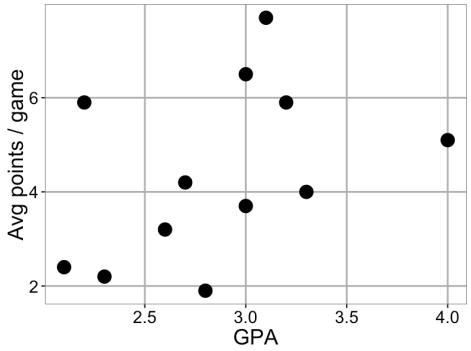


2-4. Coach Faulder's scenario shows a clear "downhill" or "negative" relationship, with a reasonable amount of strength, and a somewhat linear shape. Coach Griswold's scenario shows no real pattern. 100 m dash is clearly a better variable to pay attention to.

2-5. This is a moderately strong, positive linear association. The 140-yard player might be an outlier, but it is hard to tell with so little data. Relationship is not useful since both variables are gathered AFTER recruitment.

2-6. Answers may vary.

2-7. See scatterplot at right. There is a weak positive linear association between GPA and average points/game in this data set. If this seems surprising, remember that there is no guarantee this relationship would continue with more data points—a weak association like this can always be a trick of the data!



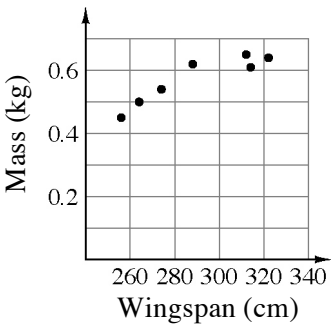
2-8. a. This is a strong, positive, nonlinear association, with two main clusters and one outlier with a height of 80 feet.

b. This is a strong, positive, linear association, with one main cluster from 1.2 to 3.1 hours and two smaller clusters. The 0.25 hours student could be an outlier.

c. This is a moderately strong, negative, linear association. No outliers or clusters.

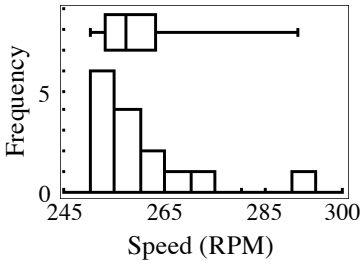
d. This is a positive, nonlinear, strong association, with one possible outlier of 16 hands, though that point may not be an outlier since it fits the pattern.

2-9. a. This is a positive, strong association that appears to be nonlinear. See graph at right. There do not appear to be any outliers.



b. A reasonable estimate based on the data might be around 0.58 kg, but confidence should be pretty low, so any numbers between 0.5 and 0.65 or so should not surprise most people; there is not enough data to make a strong estimate at this point.

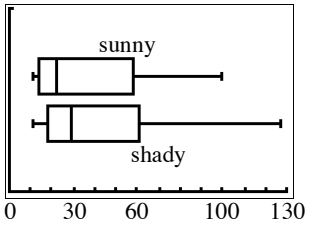
2-10. a.



b. The median is 257 RPM. The graph is single-peaked and skewed. The IQR is Q3 – Q1 = 263 – 253 = 10 RPM. 291 RPM is apparently an outlier.

c. The median. Because the data is not symmetrical and has an outlier the mean is not an appropriate measure of center.

2-11. a. See boxplots at right. Unequivocally, the farmer should plant in shade. The median crop is about 7 bushels/tree higher in shade. The minimum, maximum, first quartile, and third quartile are all higher in shade. Both distributions are skewed in the same direction. The spread in data (IQR) is almost the same for both type of tree—the middle box is the same size for both boxplots. The maximum of 127 bushels from one of the shady trees is almost certainly an outlier.



b. No. Neither of the boxplots are symmetrical; the distributions are skewed. The maximum on the shady plot may be an outlier.

