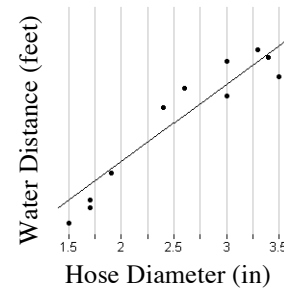# Lesson 2.2.4

2-74. a. A line may not have been the best way to model this data. The residual plot has a definite U-shaped pattern.

Water Distance (feet) vs. Hose Diameter (in)

b. The worst prediction was for a 3.5" hose. Residual was –24 feet. The model expected the water to be thrown about 24 feet more than it actually was.

c. See sample solution at right.

2-75. a. There is a positive linear association of moderate strength with no outliers.

b. The mayor said this because the more firefighters, the more damage at the fire. However, when there are more firefighters—and more damage— those fires are bigger; the firefighters are not causing the damage, the size of the blaze is. The mayor cannot assume a causation relationship.

2-76. a. There could be other variables that would cause someone who eats spinach to be physically strong.

b. Sample answer: If you are health conscious, you eat healthy things, but you also workout. It could be the working out that is causing the increased strength.

2-77. The number of hours you play in a contact sport; the more hours you play in a contact sport means more chance of a concussion, but also more hours that you wear a helmet.

2-78. a. It is possible that young people have more calcium in their diet and young people get less cancer. Also, the kind of person that has more calcium in their diet may be more health conscious, get more exercise and avoid smoking etc.

b. As times change so do many variables that may or may not be connected. The world is becoming more technologically oriented. However, fast food consumption is also booming. Perhaps it is causing fewer families to eat together and so they spend less time together. Maybe families with lots of activities use more technology and have less time to spend together so the activities are the cause.

c. It is more likely that depressed people are eating chocolate than chocolate is causing depression. Chocolate is in many foods that can cause weight gain. Perhaps weight gain is a significant cause of depression.

d. Kids who are prone to misbehavior may get spanked more than more compliant children and have a more difficult time learning and testing in school environments. Spanking may have cultural or economic links that are also connected with the amount of emphasis or value placed on education in the home.

e. People who live in poverty or are unemployed cannot afford health insurance and also cannot afford the good nutrition and healthcare which leads to less risk of dying younger. It may be that people without heath insurance (without benefits) work in more dangerous jobs.

2-79.   Answers vary.

2-80. a.   A very strong positive linear association with no outliers.

b.
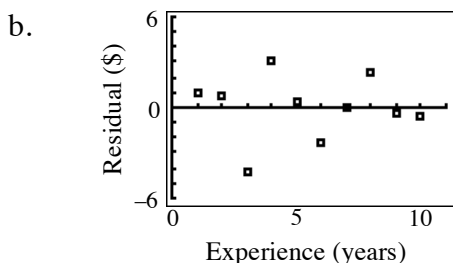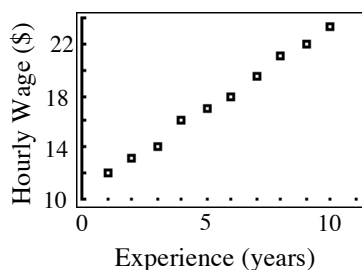
Yes, the residual plot shows random scatter with no apparent pattern.

c. $r = 0.998$, a very strong positive linear association.

2-81. a.   Each additional degree of temperature, predicts an increase of 410 park visitors.

b.   The residuals are positive, so expect the actual values to be greater than the predicted values. The predictions from the model may be too low.

c.   The residual is about –7 thousand people; the LSRL predicts 24.95 thousand people. actual – 24.95 = –7; the actual number of people in attendance was about 17,900.

d.   The predicted attendance is between 11,800 and 25,800 people.

e.   The residual plot shows a clear curve; the linear model is not appropriate. For temperatures in the 80s the model predicts too low an attendance; for temperatures in the mid 90s, the model predicts too high an attendance. The range for predicted attendance in part (d) is very large and therefore not useful.

       Statistics

2-82. a.   Choosing weight as the explanatory variable, $\widehat{speed} = 40.8 - 1.42$(weight). The slope of -1.42 means that for every additional pound a barracuda weighs its expected speed burst decreases by 1.42 mph. The 40.8 means that a barracuda that weighed nothing would expect to be able to burst at 40.8 mph, which is a nonsensical idea in this context since barracudas must have weight to exist.

       b.   The typical difference between the actual speed burst of a barracuda and the value predicted by this model is 1.8 mph.

       c.   The mean weight is 12.02 lbs so 12 should be near the middle of the data. The standard deviation of the weight is 1.85 lbs, so most of the data should be within about 3.7 on each side of the 12.02. The standard deviation of the residuals is 1.8 mph, so most of the data should be within about 3.6 of $y = 0$ on the residual plot. One reasonable set of scales is shown at right.

       d.   There is a fairly strong negative linear association between these variables. The strength and direction are confirmed by the value of $r = -0.8305$ near -1, while the linearity of the form can be seen in the randomly scattered residual plot. There are no points with extremely high residuals relative to others, so it seems there are no major outliers.
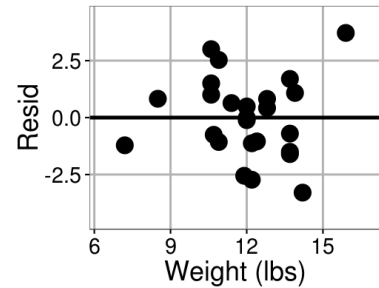
2-83. a.   Answers may vary. See sample diagram at right.

       b.   Quantitative. The data is numeric, not counts in named categories.

Mallard Lengths (cm)