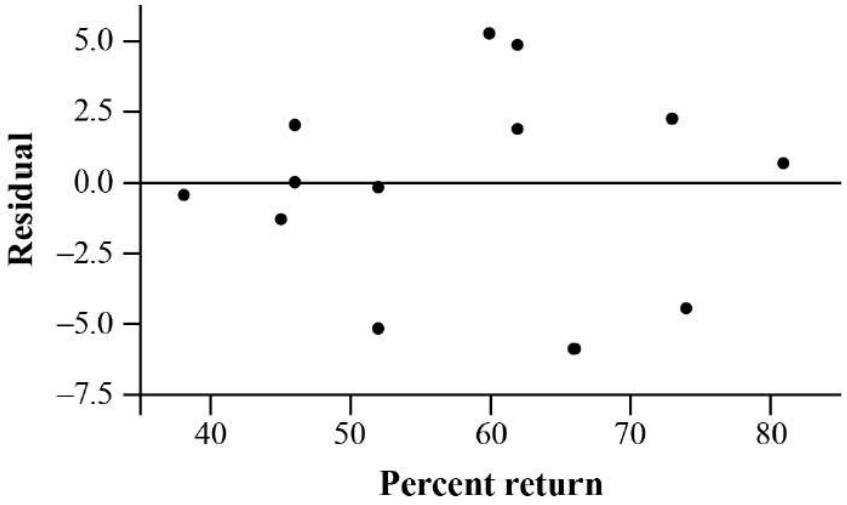
Unit 2 Homework

Assignment 3

3.49 (a) Because there is no obvious leftover pattern in the residual plot, a line is an appropriate model to use for these data.



(b) The point with the largest residual (66% returning) has a residual of about –6. This means that the colony with 66% returning birds has about 6 fewer new adults than predicted based on the percent returning.

3.51 No. Because there is an obvious negative-positive-negative pattern in the residual plot, a linear model is not appropriate for these data. A curved model would be better.

3.52 No. Because there is an obvious positive-negative-positive pattern in the residual plot, a linear model is not appropriate for these data. A curved model would be better.

3.55 (a) The predicted free skate score is $\hat{y} = -16.2 + 2.07(78.5) = 146.295$. The residual is $y - \hat{y} = 150.06 - 146.295 = 3.765$. Yu-Na Kim's free skate score was 3.765 points higher than predicted based on her short program score.

(b) Because there is no leftover pattern in the residual plot, a linear model is appropriate for these data.

(c) When using the least-squares regression line with x = short program score to predict y = free skate score, we will typically be off by about 10.2 points.

(d) About 73.6% of the variation in free skate scores is accounted for by the linear model relating free skate scores to short program scores.

3.56 (a) The predicted height is $\hat{y} = 106.1 + 4.21(10) = 148.2$. The residual is $y - \hat{y} = 141 - 148.2 = -7.2$. This student's height was 7.2 cm less than predicted based on the student's age.
(b) Because there is no leftover pattern in the residual plot, a linear model is appropriate for these data.
(c) When using the least-squares regression line with x = age to predict y = height, we will typically be off by about 8.61 cm.
(d) About 27.4% of the variation in height is accounted for by the linear model relating height to age.

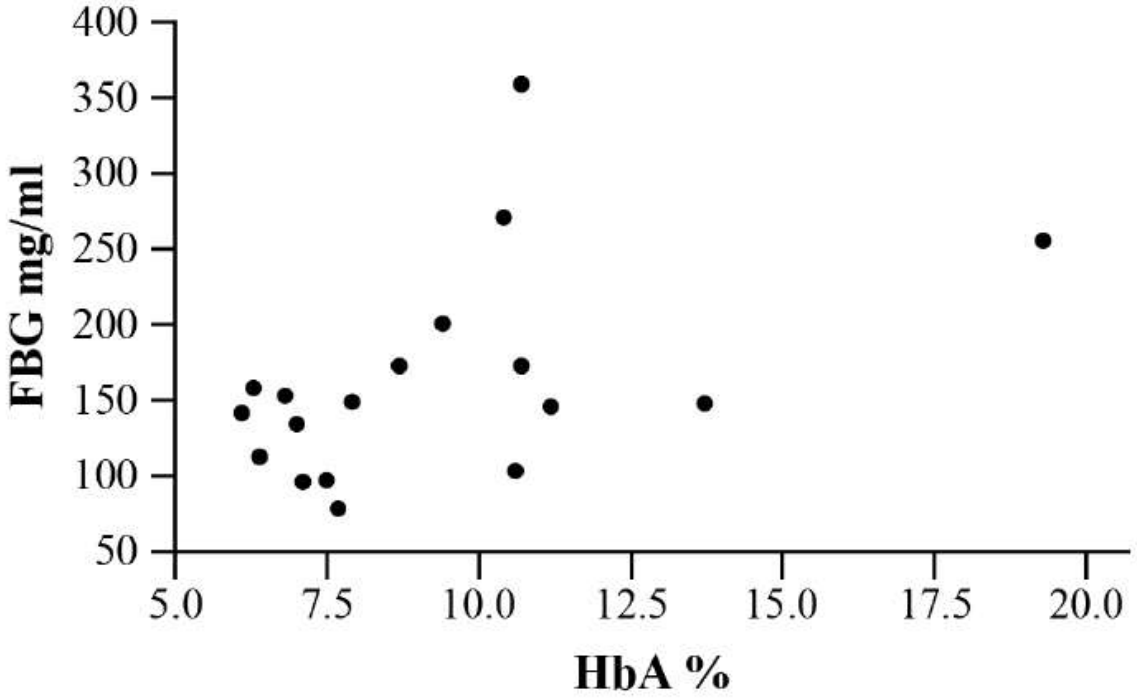
3.61 (a) The slope is $b = 0.5\left(\frac{2.7}{2.5}\right) = 0.54$. The y intercept is $a = 68.5 - 0.54(64.5) = 33.67$. So the equation for predicting y = husband's height from x = wife's height is $\hat{y} = 33.67 + 0.54x$.

(b) If the value of x is one standard deviation below $\overline{x}$, the predicted value of y will be r standard deviations of y below $\overline{y}$. So, the predicted value for the husband is $68.5 - 0.5(2.7) = 67.15$ inches.

3.64 (a) $r^2 = (0.596)^2 = 0.3552$. About 35.52% of the variation in the percent change for the entire year is accounted for by the linear model relating the percent change for the entire year to the percent change in January.

(b) When using the least-squares regression line with x = percent change in January to predict y = percent change for the entire year, we will typically be off by 8.3%.

3.69 (a) A scatterplot of this relationship is shown below. There is a moderate, positive linear association between HbA and FBG. There are possible outliers to the far right (subject 18) and near the top of the plot (subject 15).



(b) Because the point for subject 18 is in the positive, linear pattern formed by most of the data values, it will make the correlation closer to 1. Also, because the point is likely to be below the least-squares regression line, it will "pull down" the line on the right side, making the slope closer to 0. Without the outlier, the correlation decreases from $r = 0.4819$ to $r = 0.3837$ as expected. Likewise, without the outlier, the equation of the line changes from $\hat{y} = 66.4 + 10.4x$ to $\hat{y} = 52.3 + 12.1x$.

(c) The point for subject 15 makes the correlation closer to 0 because it decreases the strength of what would otherwise be a moderately strong positive association. Because this point's x coordinate is very close to $\overline{x}$, it won't influence the slope very much. However, it will make the y intercept increase because its y coordinate is so large compared to the rest of the values. Without the outlier the correlation increases from $r = 0.4819$ to $r = 0.5684$, as expected. Likewise, without the outlier, the equation of the line changes from $\hat{y} = 66.4 + 10.4x$ to $\hat{y} = 69.5 + 8.92x$.