Exercise Set #12
Answers to Odd-Numbered Exercises

12-1 (a) The response variable is \( Y = \text{“number of thousands of bacteria”} \), and the explanatory variable is \( X = \text{“temperature”} \).

(c) five-number summary for the residuals:
\[
-44.4286, -14.9018, -2.1964, +19.8036, +42.8036
\]
interquartile range for the residuals =
\[
19.8036 - (-14.9018) = 34.7054
\]

(d) There are no outliers, since
\[
-14.9018 - (-44.4286) = 29.5268 \quad \text{and} \quad 42.8036 - 19.8036 = 23.0000 \quad \text{are both less than} \quad 1.5(34.7054) = 52.0581.
\]

(e)
12-1 - continued

(f) The scatter plot constructed in part (e) appears to display a nonlinear relationship, which suggests that the residual plot will not look random.

(g) The linearity assumption does not appear to be reasonable, since the residual plot does not look random.

12-3 (a) + 5, – 9, – 36, + 15, + 5, + 42, – 8, + 17, – 14, – 17

(b) five-number summary for the residuals:

\[-36, -14, -1.5, +15, +42\]

interquartile range for the residuals \(= 15 - (-14) = 29\)
12-3 - continued
(c) There are no outliers, since \(-14 - (-36) = 22\) and \(42 - 15 = 27\) are both less than \(1.5(29) = 43.5\).

(d) The scatter plot constructed in Exercise 11-1(h) appears to display a linear relationship, which suggests that the residual plot will look random.

(e) The linearity assumption appears to be reasonable, since the residual plot looks random.

12-5 (a) It is difficult to make a decision about whether the linearity assumption is reasonable from a residual plot of this data, because there are so few data points.

(b) The linearity assumption would most likely not be reasonable, if the training hours in the data had ranged from 5 to 100 hours, since we would expect the time to perform the job to stop decreasing after a certain number of training hours.
12-7 Since the vertical axis does not begin at zero, the heights of the bars are distorted when compared with each other. For instance, the bar for “Brand X” is three times as high as the bar for “Other,” but the frequency for “Brand X” is less than two times as high as the frequency for “Other”.

12-9 In a line graph, the variable on the horizontal axis should be time, but the five different companies have been represented on the horizontal axis.