16. Baseball. An exercise in the last chapter looked at the relationship between the number of games won by American League baseball teams and the average attendance at their home games for the first half of the 2001 season. Here are the scatterplot, the residuals plot, and part of the regression analysis.

![Graphs of scatterplot, residuals, and predicted values](image)

- Dependent variable: Attendance
- $R^2 = 33.3\%$
- Variable | Coefficient | Intercept | Wins
- 5773.27 | 5271.609

a) Do you think a linear model is appropriate here? Explain.
b) Interpret the meaning of $R^2$ in this context.

d) Yes, a linear model is appropriate because the residual plot shows no obvious pattern.
b) The $R^2$ value of 33.3% demonstrates that 33.3% of the variation in the average attendance can be explained by the variation in the number of wins.

18. Second inning. Consider again the regression of average attendance on wins for the baseball teams examined in Exercise 16.

a) What is the correlation between wins and average attendance?
b) What would you predict about the average attendance for a team that is 2 standard deviations above average in games won?
c) If a team is 1 standard deviation below average in attendance, what would you predict about the number of games the team has won?

\[ r = \sqrt{r^2} = \sqrt{0.333} = 0.58 \]

Moderately strong positive correlation

b) The $S_x$ and $S_y$ are found in the slope with the $r$.

\[ r \cdot \frac{S_y}{S_x} = \frac{r \cdot S_y}{S_x} \]

If you think of it this way, then $r$ is just a scale factor. If it says 2 SDs above in games won we have:

\[ \frac{r \cdot S_y}{2S_x} \]

would have to go up .58 $\cdot$ 2 $= 1.16$ SDs.

c) Similarly, we have

\[ \frac{r}{(-1S_y)} \rightarrow \text{divide this time:} \frac{1}{0.58} = 1.74 \text{ SDs below} \]
For every additional win, there is an increase of approximately 518 people in attendance.

\[ y = 1653.72 + 517.009 \times x \]

(a) A negative residual indicates that the actual attendance was lower than the predicted attendance.

(b) \[ y = 5773.37 + 517.009 \times 50 \]

(c) The equation for the average attendance and games won by American League baseball teams is given in Exercise 16. Estimate the average attendance for a team with 50 wins.

(d) Interpret the meaning of the slope of the regression line in this context.

(e) The St. Louis Cardinals are not included in these data in this context, so their average residual means in general, what would a negative residual mean in this context?

(f) Write the equation of the regression line for the data in context.

(g) In general, what would a negative residual mean in this context?

(h) The St. Louis Cardinals are not included in this context.

(i) Interpret the meaning of the slope of the regression line in this context.

(j) The equation for the average attendance and games won by American League baseball teams is given in Exercise 16. Estimate the average attendance for a team with 50 wins.

(k) Interpret the meaning of the slope of the regression line in this context.

(l) The St. Louis Cardinals are not included in these data in this context, so their average residual means in general, what would a negative residual mean in this context?

(m) Write the equation of the regression line for the data in context.

(n) In general, what would a negative residual mean in this context?

(o) The St. Louis Cardinals are not included in this context.

(p) Interpret the meaning of the slope of the regression line in this context.

(q) The equation for the average attendance and games won by American League baseball teams is given in Exercise 16. Estimate the average attendance for a team with 50 wins.

(r) Interpret the meaning of the slope of the regression line in this context.

(s) The St. Louis Cardinals are not included in these data in this context, so their average residual means in general, what would a negative residual mean in this context?

(t) Write the equation of the regression line for the data in context.

(u) In general, what would a negative residual mean in this context?

(v) The St. Louis Cardinals are not included in this context.

(w) Interpret the meaning of the slope of the regression line in this context.

(x) The equation for the average attendance and games won by American League baseball teams is given in Exercise 16. Estimate the average attendance for a team with 50 wins.