Residuals, Residual Plots, and R^2

Homework: Pages 192-194 #1, 3, 5, 7, 11, 28

Review

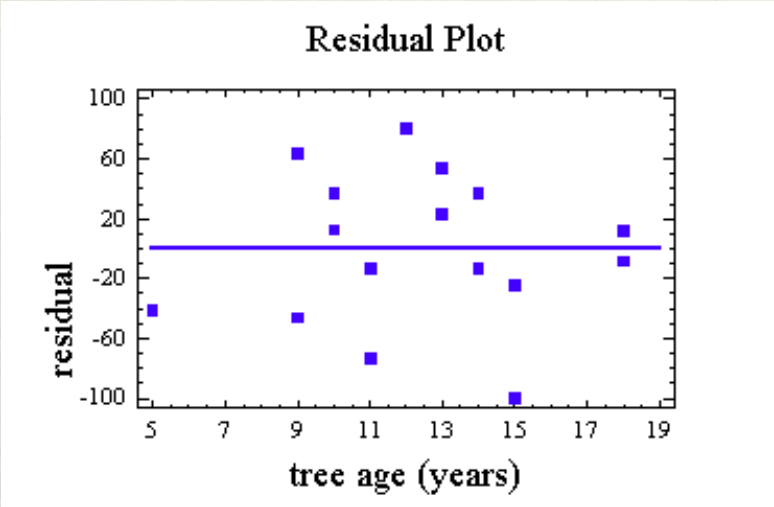
- A linear model is used to predict an individual's annual bonus based on the number of years with the company. Sandra told her husband that her residual was approximately -$862. Explain what this means.

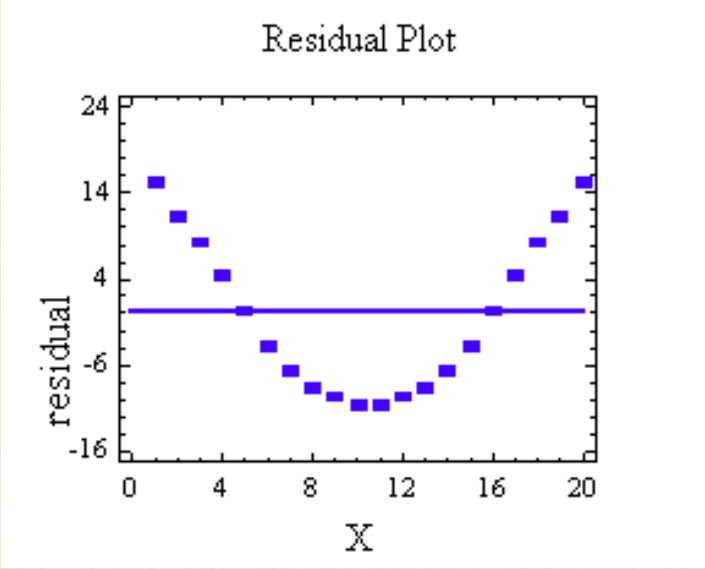
- The LSRL used to predict the bonus based on the number of years of service is $\widehat{bonus} = 3650 + 52.56(years)$. If Sandra was expecting to get a bonus of $4,500, how long has she worked with the company?

Residual Plots

- A **residual plot** is a graph that shows the **residuals** on the vertical axis and the independent variable on the horizontal axis.
- If the points in a **residual plot** are randomly dispersed around the horizontal axis, a linear regression model is appropriate for the data.
- If the residuals form a pattern, a non-linear model is more appropriate.

Examples





- Residual Plot A shows random scatter- the linear model is a good fit for the data
- Residual Plot B shows a pattern- the linear model is not a good fit for the data

Pain Relief vs. Dosage Revisited

- Recall our example of dosage and pain relief.
- If you knew nothing about regression, what would you predict the pain relief would be for an individual?

Pain Relief vs. Dosage Revisited

- What was the sum of the squared residuals from the pain relief vs. dosage example?

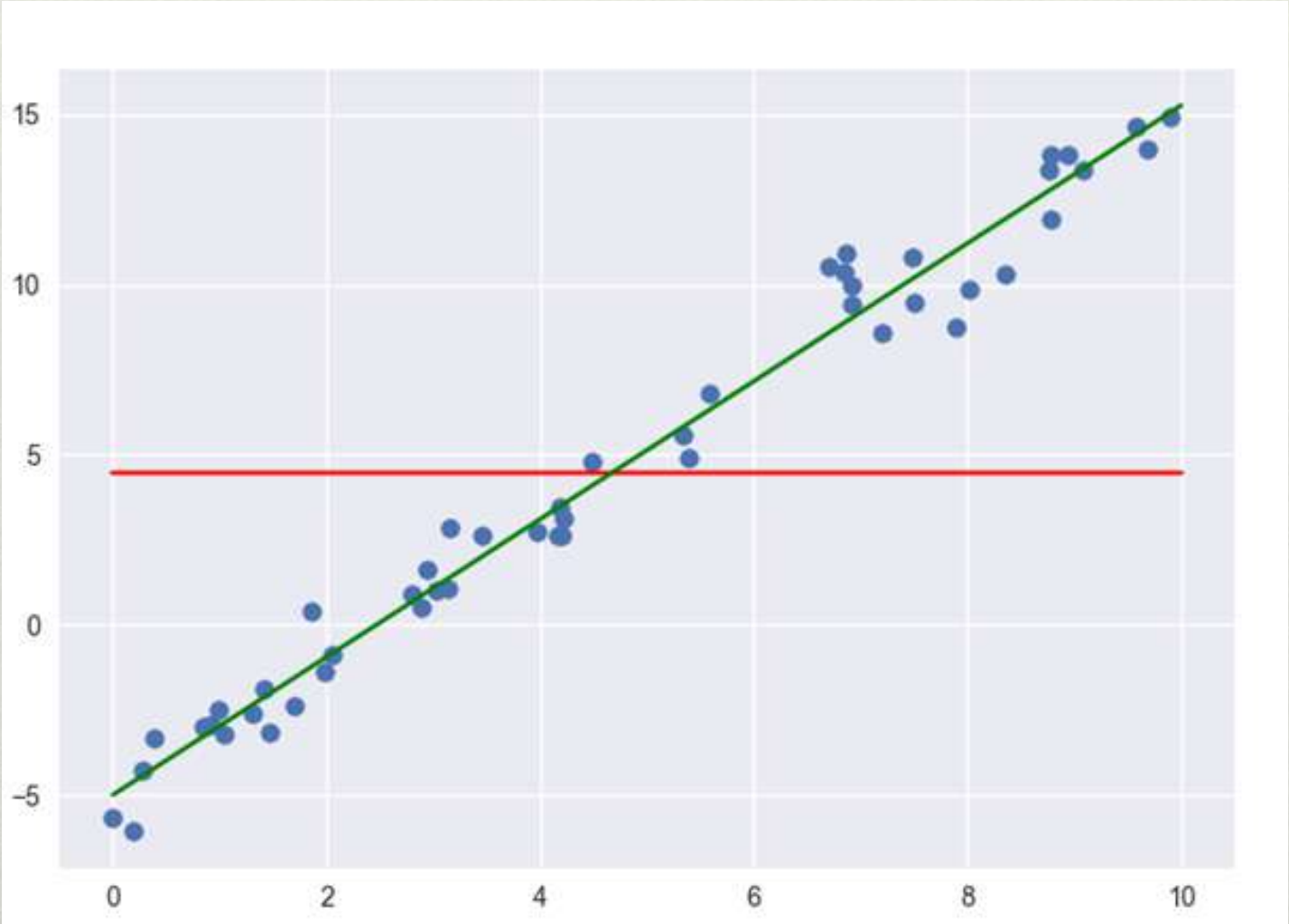
- Complete the table below:

x	y	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$	$\bar{y}$	$y - \bar{y}$	$(y - \bar{y})^2$

R^2

- Also referred to as the coefficient of determination.
- R^2 is the percent reduction of the squared residuals when comparing the regression line to a line through the mean of the y-values

R^2 Explained



Interpreting R^2

- Approximately _____ % of the variability in the y-variable can be explained by the linear regression of the y-variable on the x-variable.
- Example: Approximately 91.56% of the variability in weight can be explained by the regression of height on weight.