Chapter 8: Linear Regression

This chapter deals with linear regression. The **linear model** is just an equation of a straight line through the data.

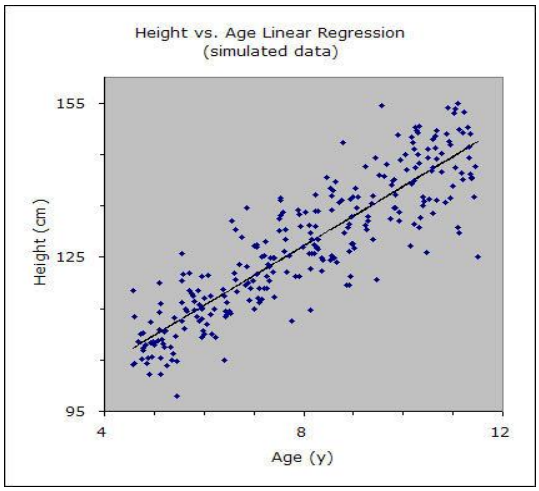
The estimates made from the linear model are called predicted values, which we label as y-hat.
The difference between the observed value and its predicted value is called the **residual**.
To find the residuals, we subtracted the predicted value from the observed.

Or

(y) – (y-hat) = residual.

The **line of best fit** is the line for which the sum of the squared residuals is smallest.

Here's an example: This picture below shows the relationship between Height and Age. By subtracting the observed values minus the expected, we can find out the residuals.



The equation for the regression line of best fit is:

$$\hat{y} = b_0 + b_1 x$$

Where b_0 is the intercept, or $b_0 = y - b_1 x$

and

b_1 is the slope, or $b_1 = r s_y / s_x$

When analyzing the line of best fit, you'll be given R^2, which is the square of the correlation between y and x.