Regression

The **regression line** models the linear relationship between two quantitative variables.

It minimizes the sum of the squared vertical distances of the points to the line.

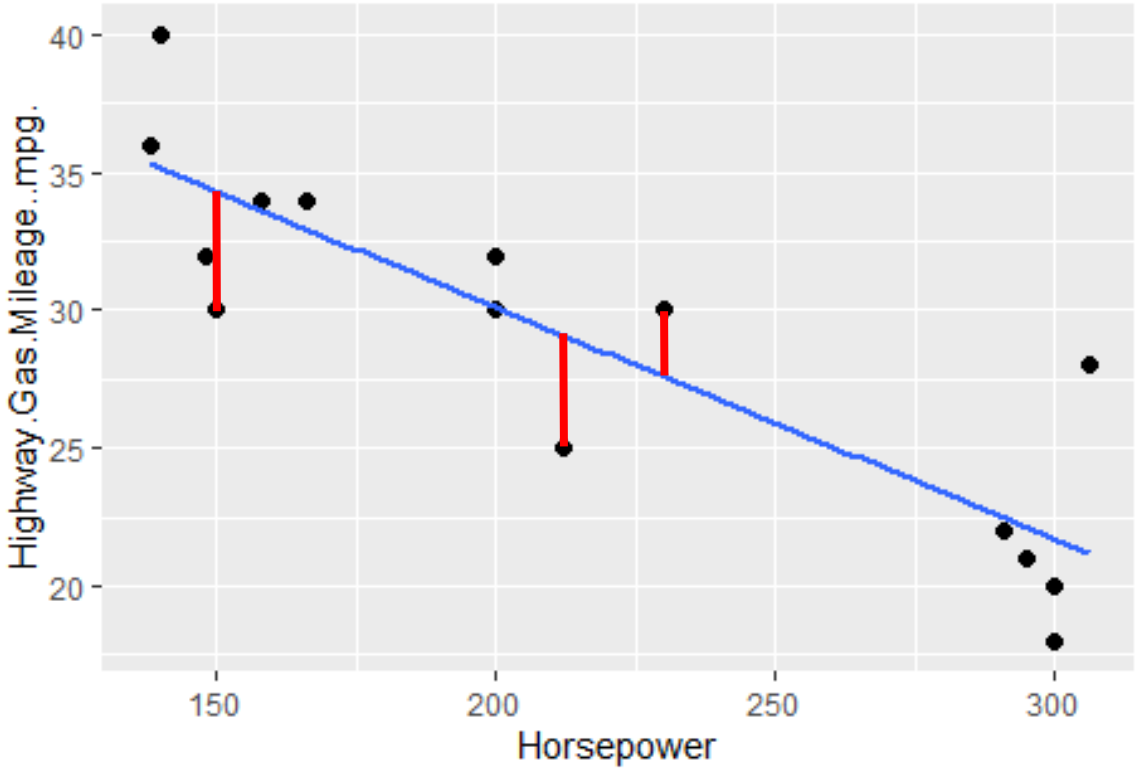
$$\hat{y} = -.084x + 46.9$$

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

y-intercept slope

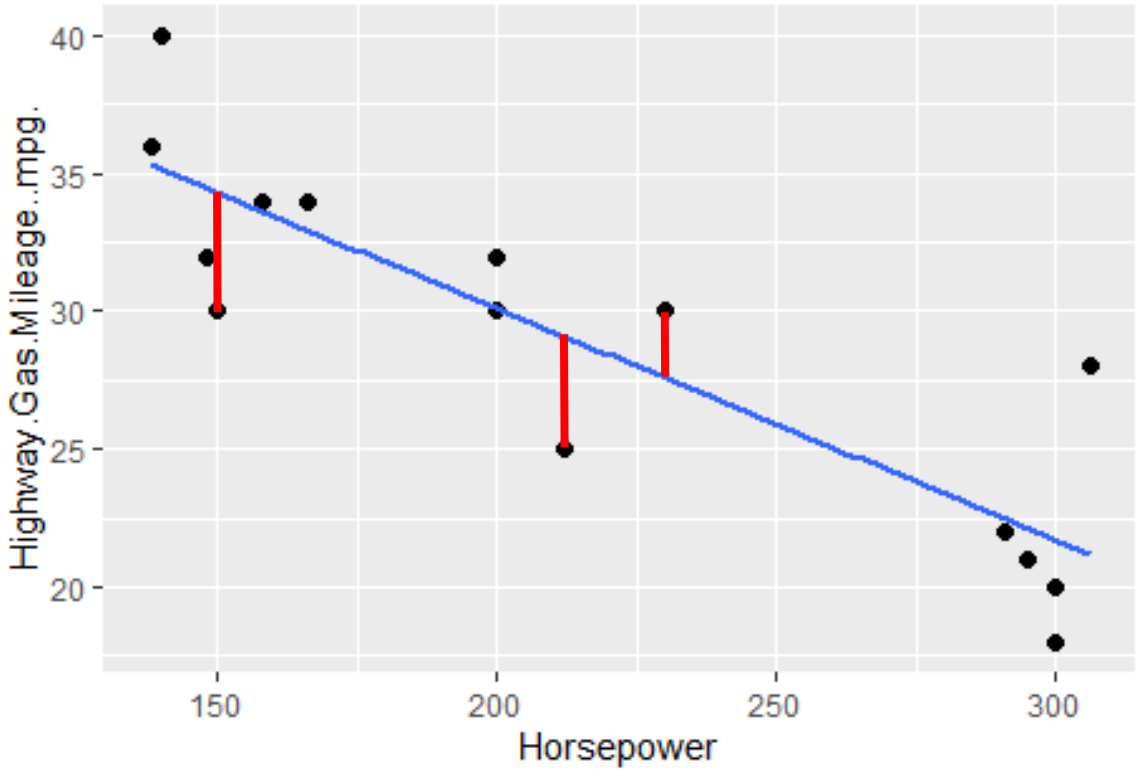
$$y - \bar{y} = \frac{r s_y}{s_x}(x - \bar{x})$$

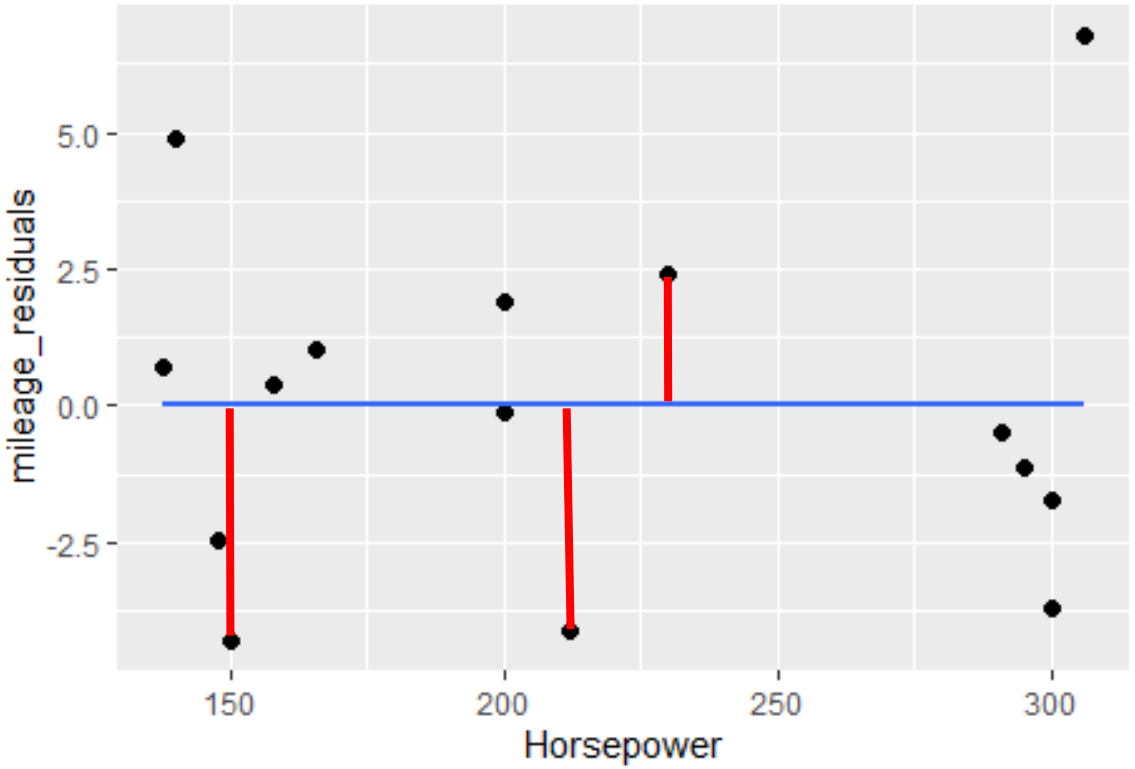
mean, standard deviation, correlation



Residuals

The **residuals** are the vertical distances of the points to the line.
They measure the error of the regression line in predicting the *y* variable.



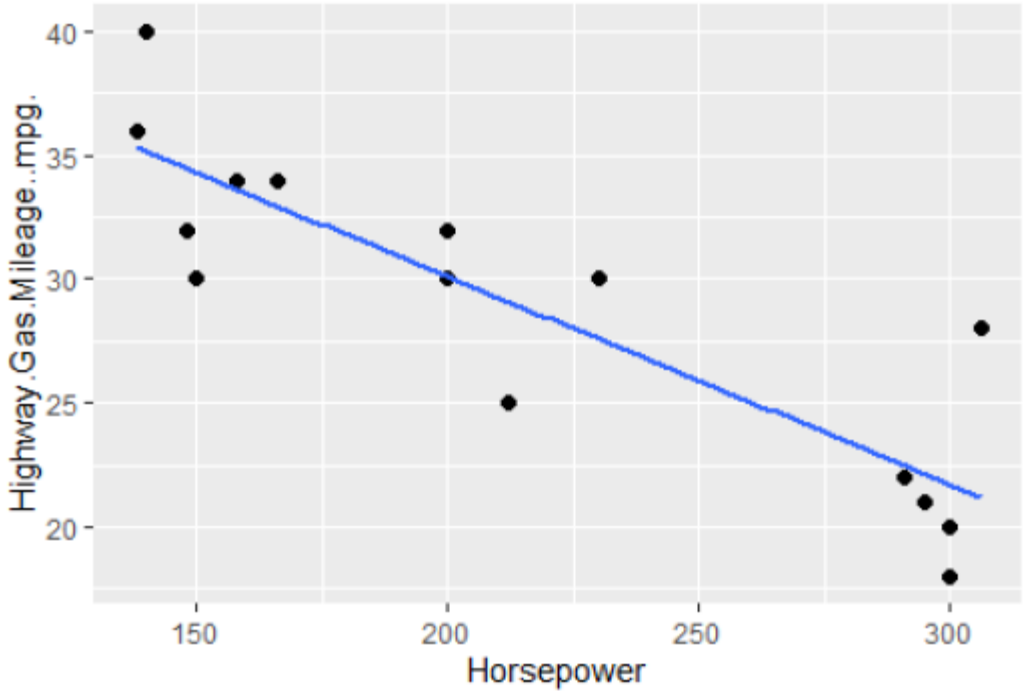


The squared correlation R^2

The number R^2 measures the proportion of the variance of the *y* variable predicted by the *x* variable.

$r = -0.8686827$

$r^2 = R^2 = 0.7546096$



1