Multiple Regression

Stat 230

## Example: Abalone age

Abalone are a type of mollusk that are harvested for their meat. Knowing the age of abalone can help set sustainable fishing quotas; however, the process is time consuming and labor intensive. Researchers collected various physical measurements, such as length, diameter, height, and weight, that are easy to collect to develop a method to predict abalone age. Below is R output for a multiple linear regression model predicting rings (a proxy for age) that uses length (mm), diameter (mm), height (mm), and whole\_weight (in grams) as predictor variables.

Call:  
lm(formula = rings ~ length + diameter + height + whole\_weight,   
 data = abalone)  
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 2.90861 0.30922 9.406 < 2e-16  
length -12.04857 2.10217 -5.731 1.07e-08  
diameter 25.64381 2.57384 9.963 < 2e-16  
height 20.24276 1.78180 11.361 < 2e-16  
whole\_weight 0.06589 0.22557 0.292 0.77   
---  
  
Residual standard error: 2.59 on 4172 degrees of freedom  
Multiple R-squared: 0.3556, Adjusted R-squared: 0.3549   
F-statistic: 575.4 on 4 and 4172 DF, p-value: < 2.2e-16

1. Report the estimated regression equation.
2. Using the model, predict the number of rings for an abalone with length 0.5 mm, diameter 0.4 mm, height 0.1 mm, and whole weight of 1 gram.
3. Interpret the slope coefficient for length in context.
4. State the hypotheses for the test of the slope coefficient for whole\_weight. What do you conclude based on the results of this test?
5. Calculate a 95% confidence interval for the slope coefficient for length. Interpret this interval in context.
6. What proportion of the variability in rings is explained by the model?