***Chapter 6: Multiple Linear Regression***

A basketball coach is curious as to what factors are associated with his player’s success at the free throw line. He uses a multiple linear regression model of *FTPercent* (a player’s free throw percentage) on *Guard* (1 for guards, 0 for other players) and *FGPercent* (a player’s shooting percentage on non free-throws). Here’s the resulting output, using data from the 2016-17 season.

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.57428 0.26671 2.153 0.0635 .

FTPercent 0.21302 0.56086 0.380 0.7140

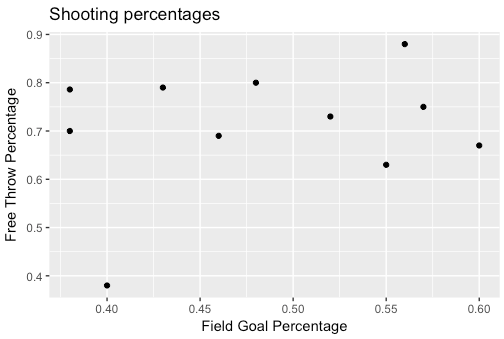
Guard 0.07351 0.08623 0.852 0.4188

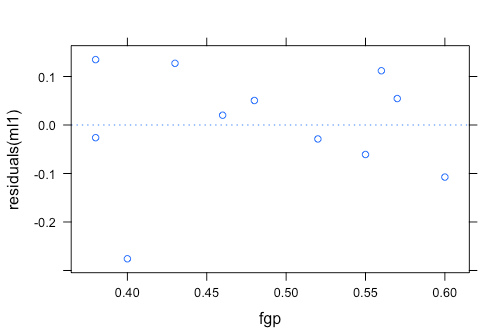
Residual standard error: 0.1367 on 8 degrees of freedom

Multiple R-squared: 0.1248, Adjusted R-squared: -0.09395

F-statistic: 0.5706 on 2 and 8 DF, p-value: 0.5866

1. What are the explanatory variables, and what is the response?
2. Write the estimated regression line for all players
3. Write the estimated regression line for any players who are Guards (*Guards* = 1)
4. Write the estimated regression line for any players who are not Guards (*Guards* = 0)
5. Here’s a plot of the team’s percentages. What does this plot tell you about your assumptions for a linear model?



1. Predict the free throw percentage for Back Jyrne, a player whose *FGPercent*  is .382% and who plays the guard position.
2. Back’s observed free throw percentage is .786. Calculate his residual, and identify how Back performed relative to his shooting percentage and his position.
3. Is either explanatory variable significantly associated with *FTPercent*? How can you tell?
4. What is the R2 and the R2adj. Why did you know the R2adj would be lower? Interpret the R2 value.
5. ****Here are two residual plots. What are they telling us about our model?

