Assignment 5

Chase Baggett
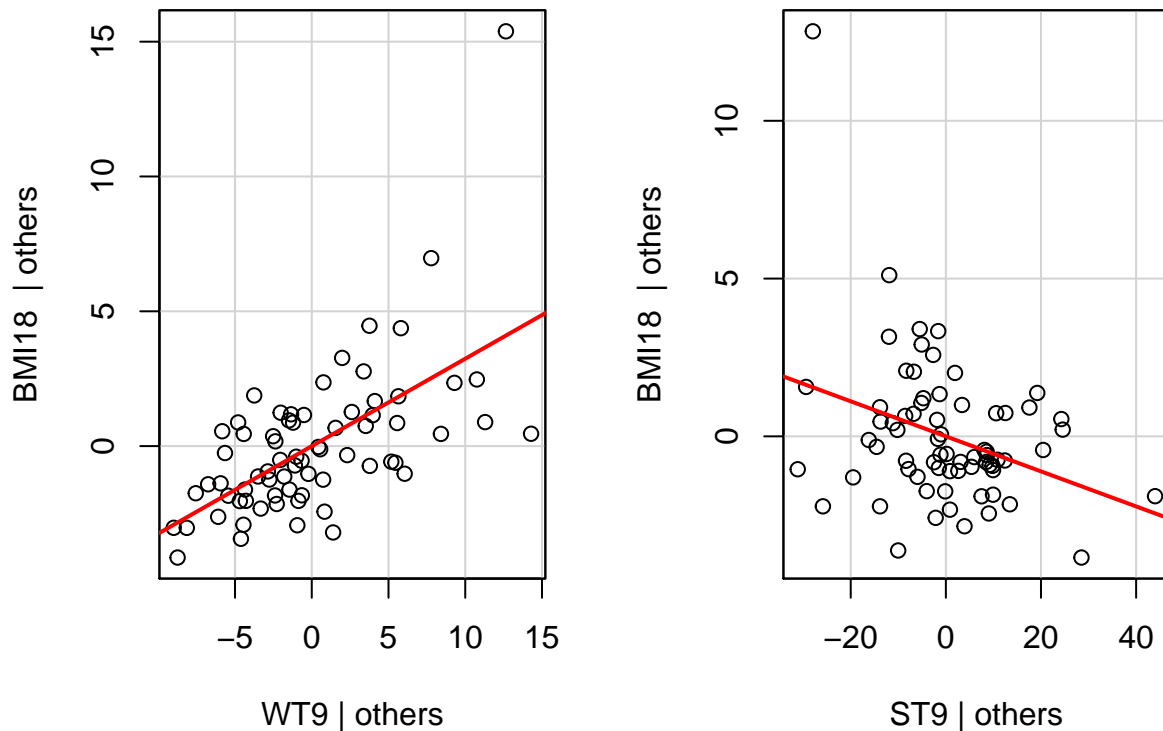
October 19, 2017

5.1

5.1.1

```
initial_model <- with(BGSgirls,lm(BMI18~WT9))
second_model <- with(BGSgirls,lm(BMI18~WT9 + ST9))
residual_model <- lm(resid(initial_model)~resid(second_model))
avPlots(second_model,main="Added Varialbe Plots of Both WT9 and ST9")
```


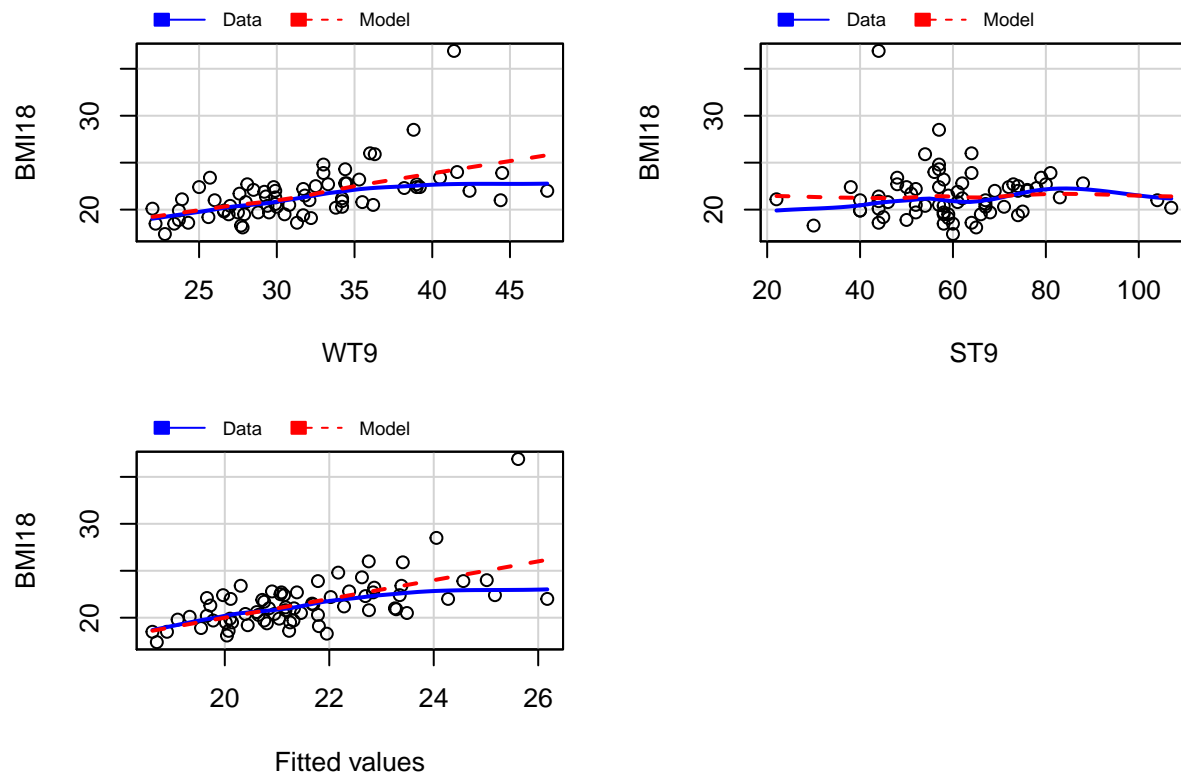Added Varialbe Plots of Both WT9 and ST9

```
ggplot(BGSgirls,aes(x=WT9,y=ST9)) + geom_point() + geom_smooth() + ggtitle("Plot of WT9 vs ST9")

## `geom_smooth()` using method = 'loess'
```

Plot of WT9 vs ST9
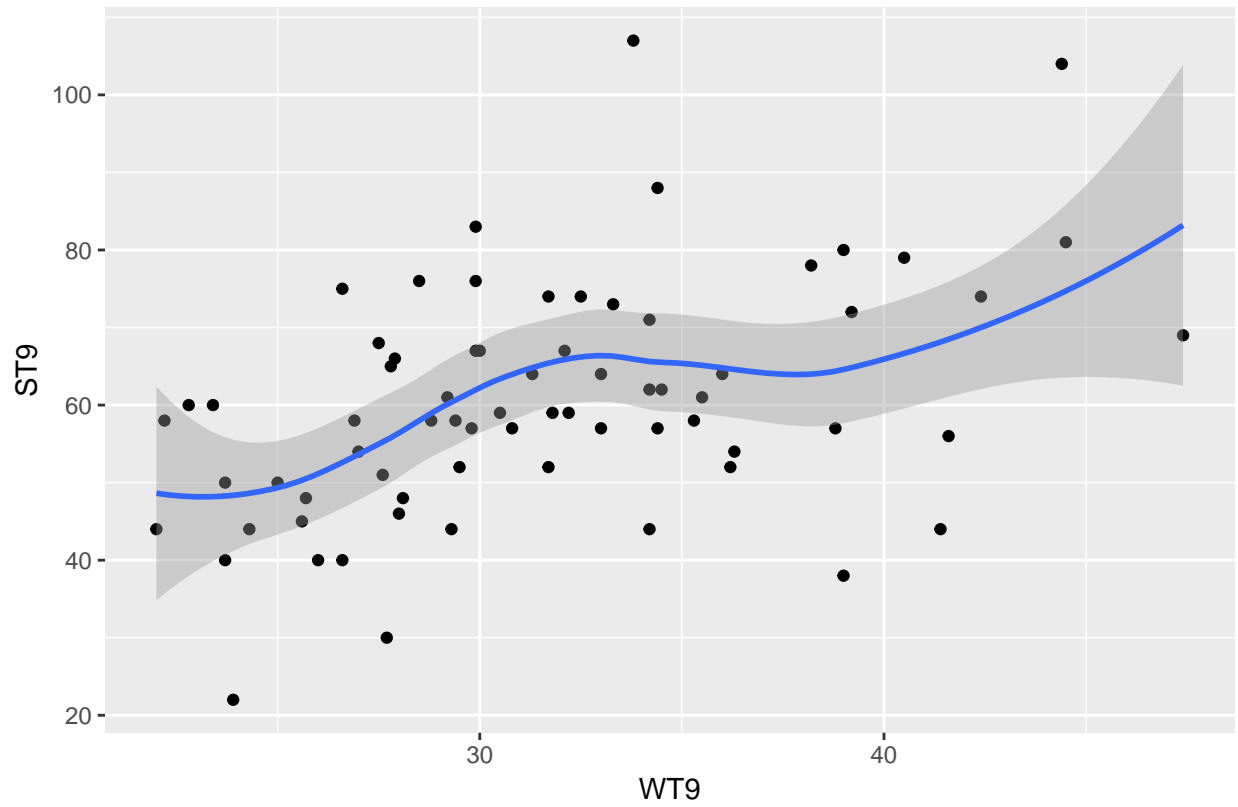
```
mmps(second_model,layout = c(2, 2))
```

Marginal Model Plots

The added variable plots let us look at the effect of one variable after adjustment for others. So in this added variable plot we can see the effect of ST9 and WT9 on BMI18 after the other adjustment for other predictor. We can see that ST9 is negatively correlated after adjustment for WT9, and WT9 is positively correlated after adjustment for ST9.

From looking at the marginal plots we can see that our model's fit declines as we attempt to predict higher fitted values.

5.1.2

```
model <- with(BGSgirls,lm(BMI18~HT2 + WT2 + HT9 + WT9 + ST9))
model_summary <- summary(model)
```

R-Squared

```
model_summary$r.squared
```

```
## [1] 0.4431186
```

Sigma

```
model_summary$sigma
```

```
## [1] 2.139798
```

T-Values

```
coef(model_summary)[, "t value"]
```

```
## (Intercept)          HT2          WT2          HT9          WT9          ST9
##  3.51381240 -1.48294722 -1.14007251  0.08362855  5.58110219 -1.99905072
```

Hypothesis Test and Conclusion We are testing the hypothesis that our coefficients are equal to zero with a two sided test. With the exception of our intercept and ST9, we fail to prove significance.

5.2.3

I am assuming X to refer to the n x (p+1) matrix of X values as notated as such in the book on page 60. Therefore, the first row of X would be the first row of the matrix of predictors. I pull the matrix out of the model, pull the first row, remove the first column which is the Y.
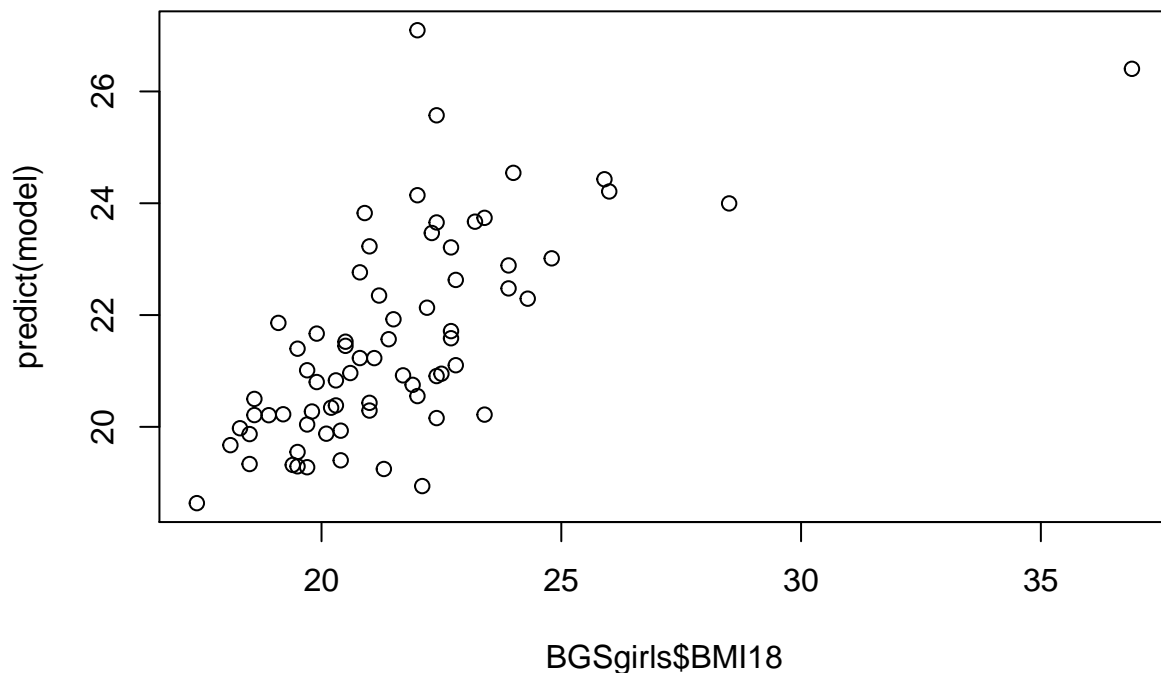
```
model$model[1,-1]
```

```
##     HT2  WT2   HT9  WT9 ST9
## 1 87.7 13.6 133.4 32.5  74
```

5.1.4

A

R^2 is the ratio of the change in RSS from one model to another. The response vs fitted values shows relationship to the RSS of a single model.

```
plot(BGSgirls$BMI18,predict(model))
```

5.1.5

If we take the square root of the values of the variance-covariance matrix, we get the standard errors on the diagonals, which we can see matches the standard errors calculated by the model. The fourth diagnal element is the standard error for HT9, 0.09634358. ###B

```
coef(model_summary)[, "Std. Error"]
```

```
## (Intercept)         HT2         WT2         HT9         WT9         ST9
##  8.78115599  0.13081860  0.27873612  0.09634358  0.07521138  0.02221856
```

```
sqrt(abs(vcov(model)))
```

```
##               (Intercept)        HT2        WT2        HT9        WT9
## (Intercept)     8.7811560 0.58420014 0.42653127 0.67328755 0.50737200
## HT2             0.5842001 0.13081860 0.12953203 0.08853074 0.05030915
## WT2             0.4265313 0.12953203 0.27873612 0.07263836 0.10379303
## HT9             0.6732876 0.08853074 0.07263836 0.09634358 0.06258317
## WT9             0.5073720 0.05030915 0.10379303 0.06258317 0.07521138
## ST9             0.2411423 0.02639219 0.03770174 0.03170920 0.01295268
##                     ST9
## (Intercept) 0.24114226
## HT2         0.02639219
## WT2         0.03770174
## HT9         0.03170920
## WT9         0.01295268
## ST9         0.02221856
```

5,1,5

```
deg_freedom <- nrow(model$model)
qchisq(.95,df=deg_freedom)
```
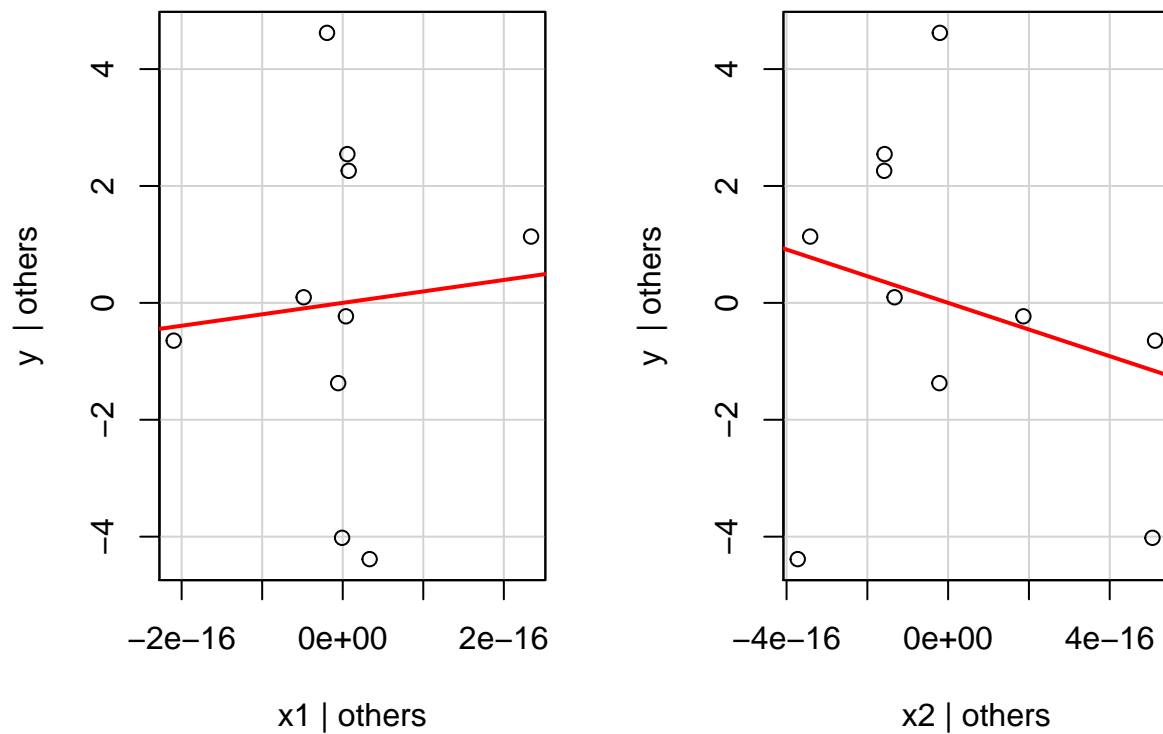
```
## [1] 90.53123
```

5.2

3.4.1

I used a random normal dataset to test this. The line plot appears to show a line that has roughly 2.2 times the slope of the added variable plot for x1.

```
data <- data.frame(y=1:10)
data$x1 <- rnorm(10)
data$x2 <- data$x1*2.2
model <- with(data,lm(y ~ x1 + x2))
avPlots(model)
```
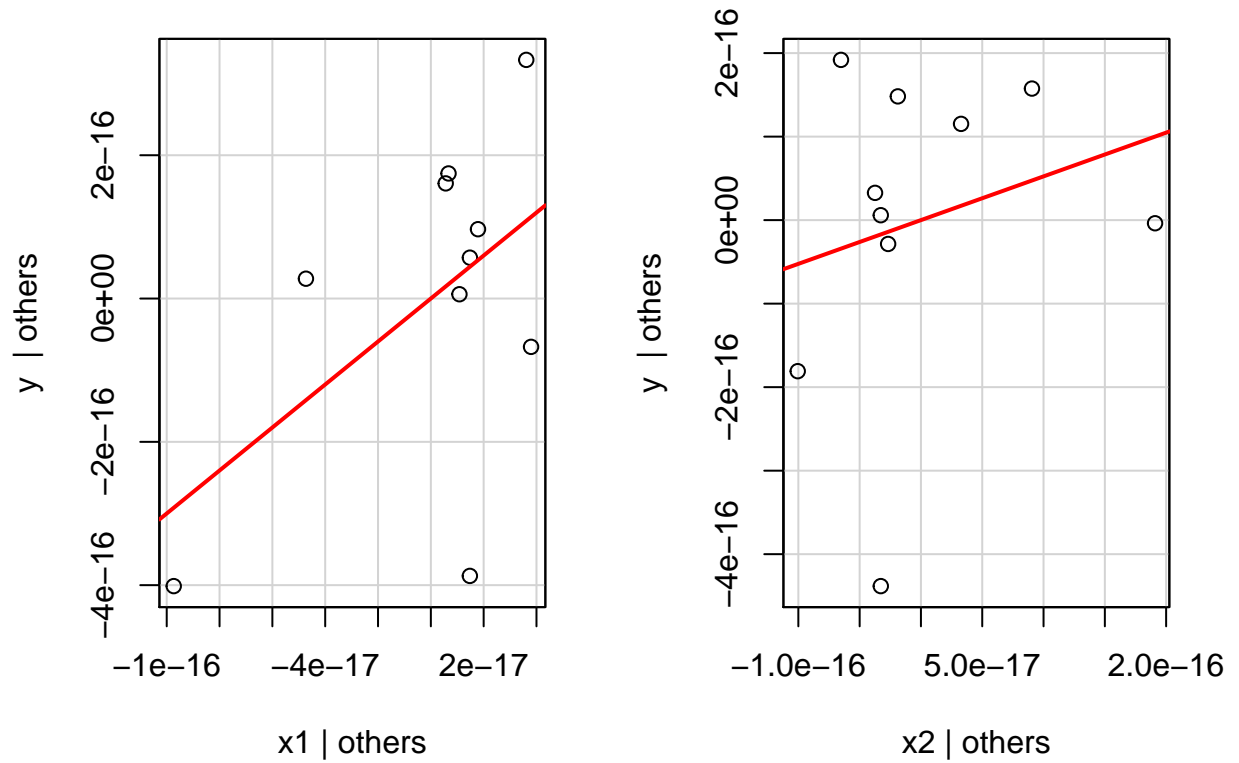
Added−Variable Plots

3.4.2

In this situation X1 is a perfect predictor of y, therefore it can't improve the model.

```
data <- data.frame(y=1:10)
data$x1 <- rnorm(10)
data$x2 <- data$x1*2.2
data$y <- data$x1 * 3
model <- with(data,lm(y ~ x1 + x2))
avPlots(model)
```

Added−Variable Plots

3.4.3

When the additional information gained by adding X2 is completely unique to x2, that is to say that none of the variation had already been explained by x1.

3.4.4

True. Additional predictors can't cause a model to decline because of the option to reduce the coefficient to zero, given in least squares we are optimizing for this vertical distance.