Lab4

Clayton Glenn

February 12, 2018

Task 1

Get Working Directory

```
getwd()
```

```
## [1] "C:/Users/cglen/Documents/Stat Methods/Labs/LAB4"
```

Task 2

Read Spruce Data File and Show Tail of Data

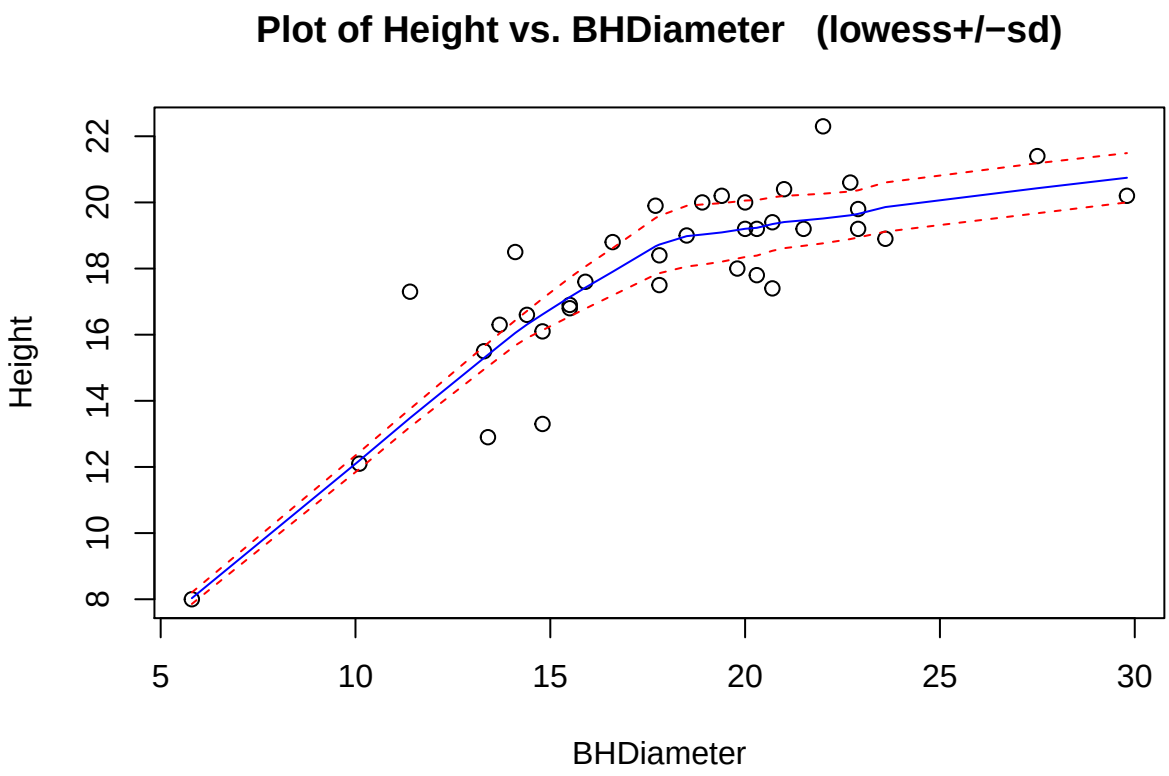
```
spruce.df <- read.csv("SPRUCE.csv", header=TRUE)
tail(spruce.df)
```

```
##    BHDiameter Height
## 31       17.7   19.9
## 32       20.7   19.4
## 33       21.0   20.4
## 34       13.3   15.5
## 35       15.9   17.6
## 36       22.9   19.2
```

Task 3

Trend Scatter of Height V BHD

```
library(s20x)
trendscatter(Height ~ BHDiameter, f = 0.5, data = spruce.df)
```

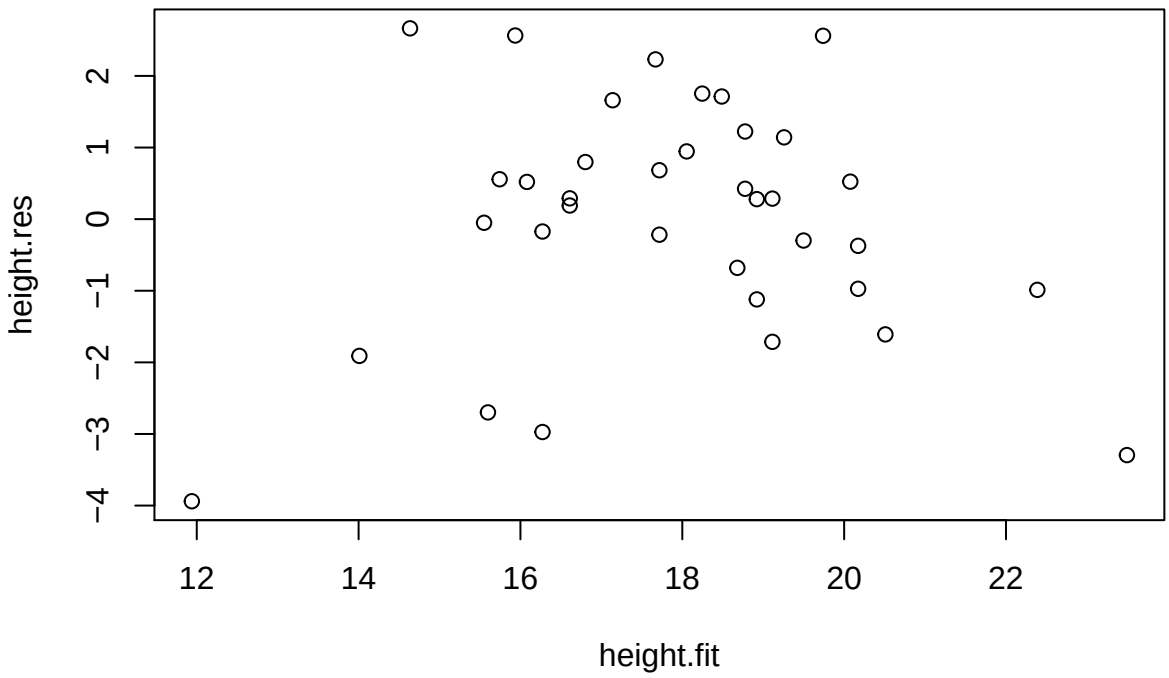


Store Linear Model, Res, and Fit

```
spruce.lm <- lm(Height ~ BHDiameter, data = spruce.df)
height.res <- residuals(spruce.lm)
height.fit <- fitted(spruce.lm)
```

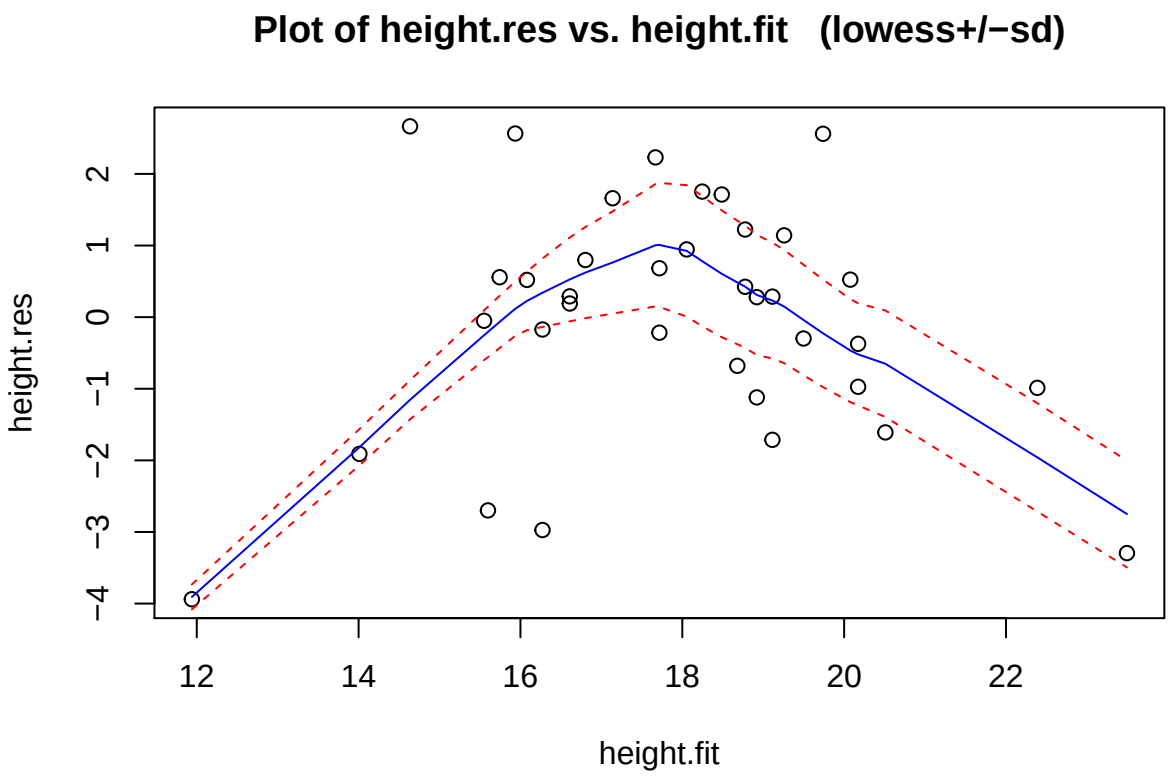
Plot of Res v Fit

```
plot(height.res ~ height.fit)
```



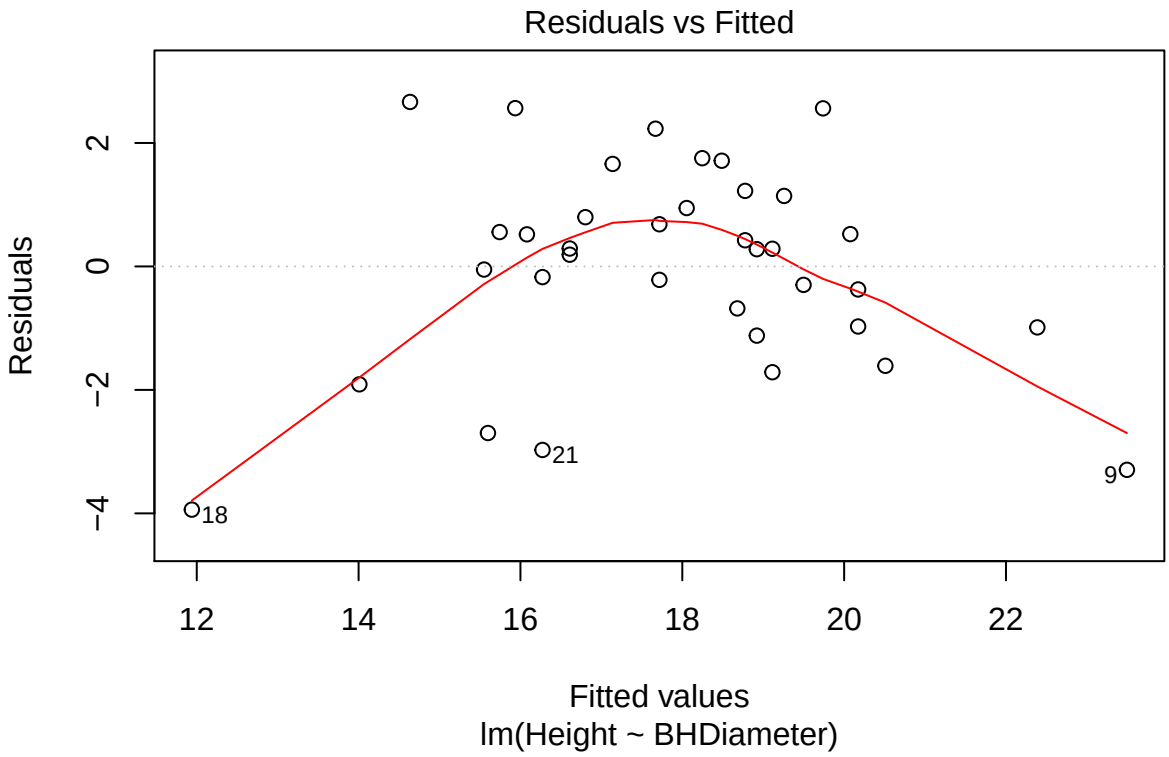
Trend Scatter of Res v Fit

```
library(s20x)
trendscatter(height.res ~ height.fit)
```



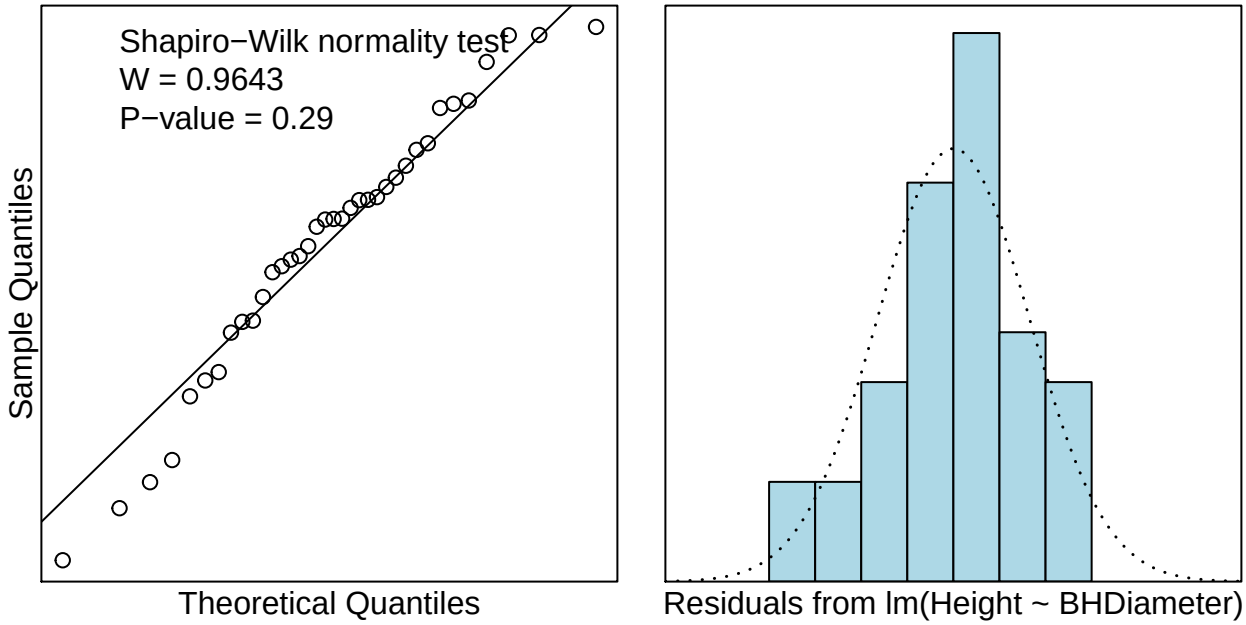
Plot Residuals V Fitted with LM

```
plot(spruce.lm, which =1)
```



Norm Check of Linear Model with Shapiro Walk

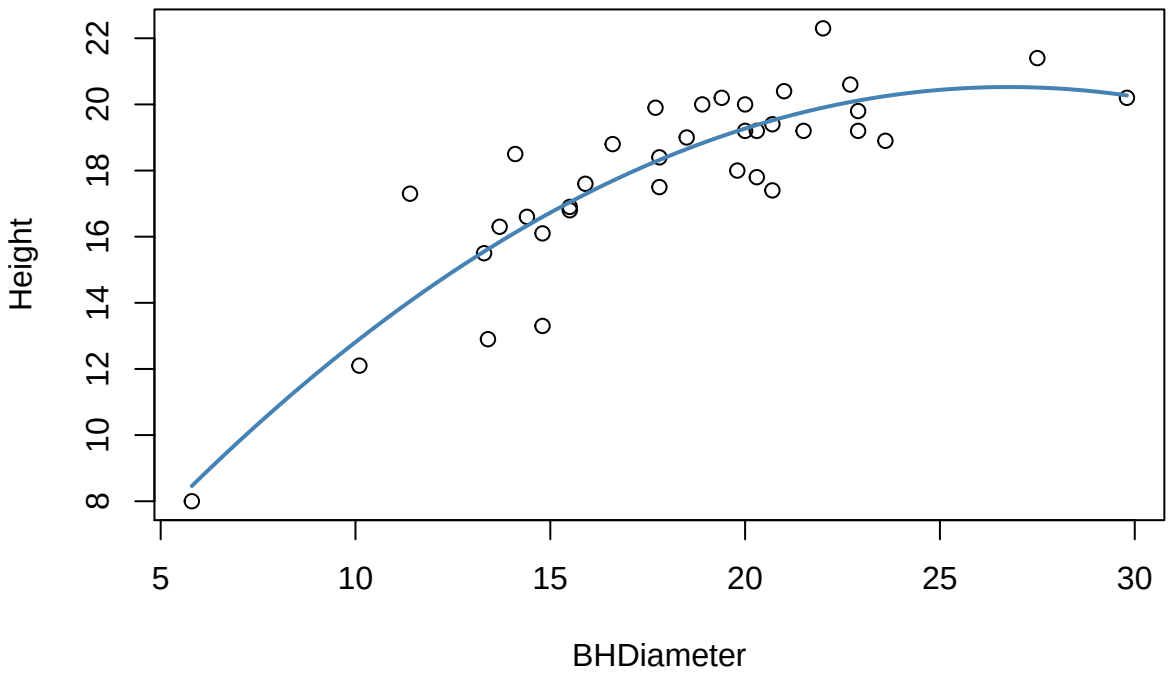
```
library(s20x)
normcheck(spruce.lm,shapiro.wilk = TRUE)
```



Task 4

Create Height v BHD Plot with new Quad Curve

```
quad.lm=lm(Height~BHDiameter + I(BHDiameter^2),data=spruce.df)
myplot=function(x){
  quad.lm$coef[1] +quad.lm$coef[2]*x  + quad.lm$coef[3]*x^2
}
plot(Height~BHDiameter, data = spruce.df)
curve(myplot, lwd=2, col="steelblue",add=TRUE)
```

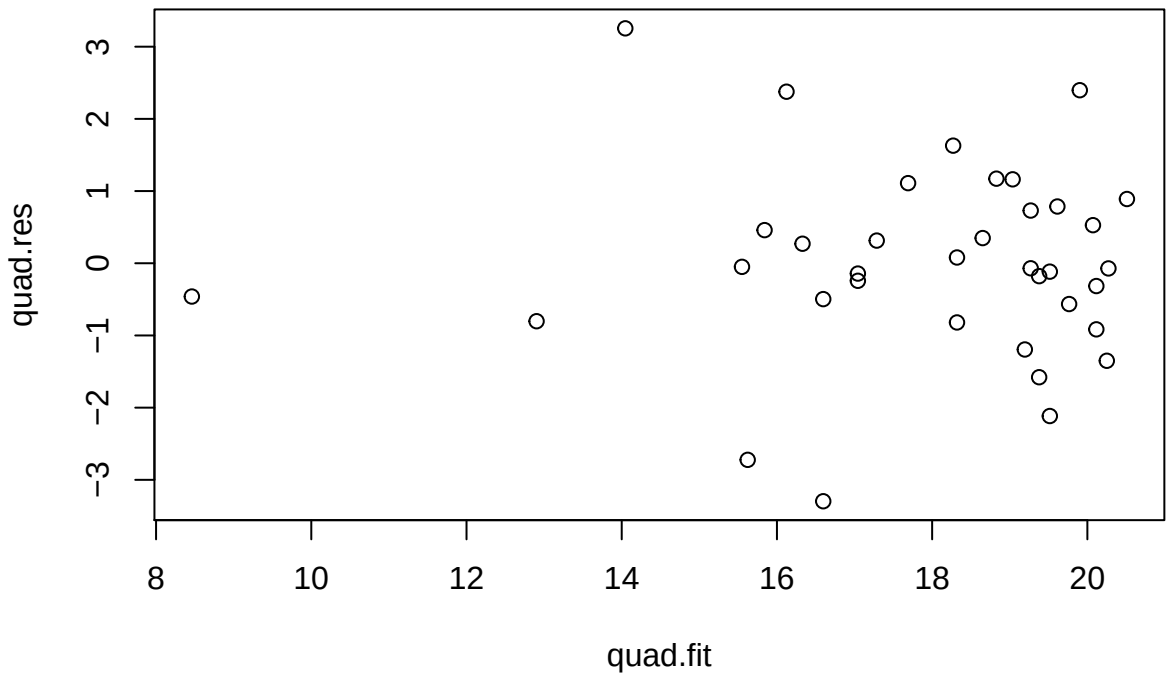


Storing Res and Fit Variables

```
quad.fit <- fitted(quad.lm)
quad.res <- residuals(quad.lm)
```

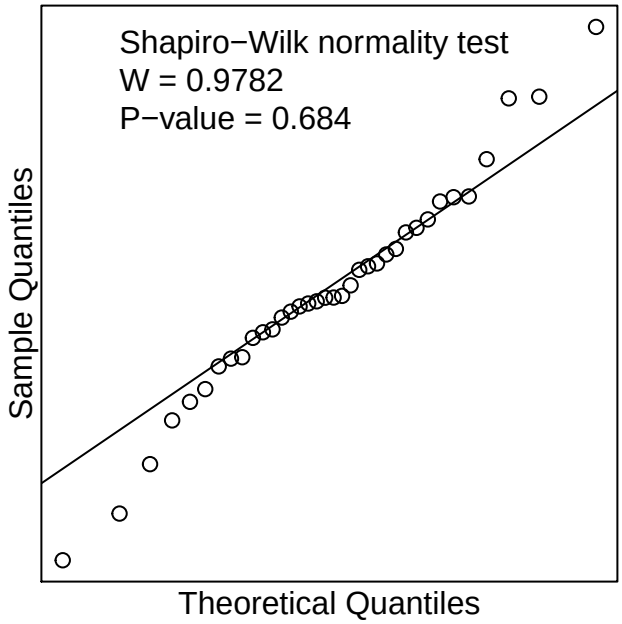
Plot with Quadratic Residuals v Fitted Values

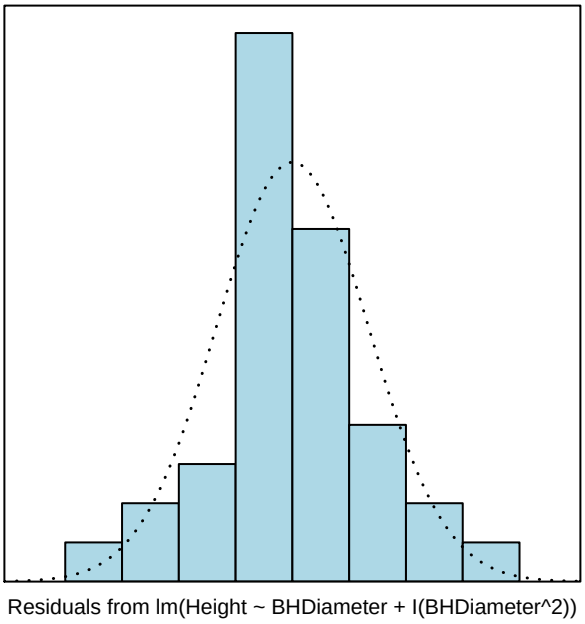
```
plot(quad.res ~ quad.fit)
```



Quadratic Norm Check and Shapiro

```
normcheck(quad.lm,shapiro.wilk = TRUE)
```





Task 5

Summary of Quadratic Model

```
summary(quad.lm)
```

```
##
## Call:
## lm(formula = Height ~ BHDiameter + I(BHDiameter^2), data = spruce.df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2966 -0.6245 -0.0707  0.7442  3.2541
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.860896   2.205022   0.390 0.698731
## BHDiameter       1.469592   0.243786   6.028 8.88e-07 ***
## I(BHDiameter^2) -0.027457   0.006635  -4.138 0.000227 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.382 on 33 degrees of freedom
## Multiple R-squared:  0.7741, Adjusted R-squared:  0.7604
```

```
## F-statistic: 56.55 on 2 and 33 DF,  p-value: 2.182e-11
```

Coefficients of Quad.lm

```
quad.lm$coef[1]
```

```
## (Intercept)
##   0.8608958
```

```
quad.lm$coef[2]
```

```
## BHDiameter
##   1.469592
```

```
quad.lm$coef[3]
```

```
## I(BHDiameter^2)
##     -0.02745726
```

Predict 15,18,and 20 using Linear Model

```
predict(spruce.lm, data.frame(BHDiameter=c(15,18,20)))
```

```
##        1        2        3
## 16.36895 17.81338 18.77632
```

Predict 15,18,and 20 using Quadratic Model

```
predict(quad.lm, data.frame(BHDiameter=c(15,18,20)))
```

```
##        1        2        3
## 16.72690 18.41740 19.26984
```

Prediction of all Spruce Heights with Quad.lm

```
with(spruce.df,predict(quad.lm,data.frame(BHDiameter)))
```

```
##         1         2         3         4         5         6         7
## 18.828182 17.690005 17.042969 17.042969 19.037171 15.840856 19.269837
##         8         9        10        11        12        13        14
## 20.510131 20.271601 19.378756 19.194478 20.115697 19.378756 16.123368
##        15        16        17        18        19        20        21
## 19.269837 12.902862 19.902612  8.460868 20.250678 19.516294 16.596623
##        22        23        24        25        26        27        28
## 18.320080 20.072189 14.045902 18.651105 16.329487 19.765011 15.623206
##        29        30        31        32        33        34        35
## 16.596623 18.320080 18.270594 19.516294 19.613682 15.549558 17.285943
##        36
## 20.115697
```

Multiple R^2 Quadratic Fitted Values

```
(with(spruce.df, sum((Height-mean(Height))^2))-with(spruce.df, sum((Height-quad.fit)^2)))/with(spruce.d
```

```
## [1] 0.7741266
```

Multiple R^2 Linear Fitted Values

```
(with(spruce.df, sum((Height-mean(Height))^2))-with(spruce.df, sum((Height-height.fit)^2)))/with(spruce
```

```
## [1] 0.6569146
```

Compare Quad.lm and Spruce.lm with Anova

```
anova(spruce.lm,quad.lm)
```

```
## Analysis of Variance Table
##
## Model 1: Height ~ BHDiameter
## Model 2: Height ~ BHDiameter + I(BHDiameter^2)
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     34 95.703
## 2     33 63.007  1    32.696 17.125 0.0002269 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(quad.lm)
```

```
## Analysis of Variance Table
##
## Response: Height
##                 Df  Sum Sq Mean Sq F value    Pr(>F)
## BHDiameter       1 183.245 183.245  95.975 2.701e-11 ***
## I(BHDiameter^2)  1  32.696  32.696  17.125 0.0002269 ***
## Residuals       33  63.007   1.909
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(spruce.lm)
```

```
## Analysis of Variance Table
##
## Response: Height
##             Df  Sum Sq Mean Sq F value    Pr(>F)
## BHDiameter   1 183.245 183.245  65.101 2.089e-09 ***
## Residuals   34  95.703   2.815
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

RSS, MSS, TSS, and MSS/TSS Calculations

```
RSS <- with(spruce.df, sum((Height-quad.fit)^2))
MSS <- with(spruce.df, sum((quad.fit-mean(Height))^2))
TSS <- with(spruce.df, sum((Height-mean(Height))^2))
RSS
```

```
## [1] 63.00683
```

```
MSS
```

```
## [1] 215.9407
```

```
TSS
```

```
## [1] 278.9475
```

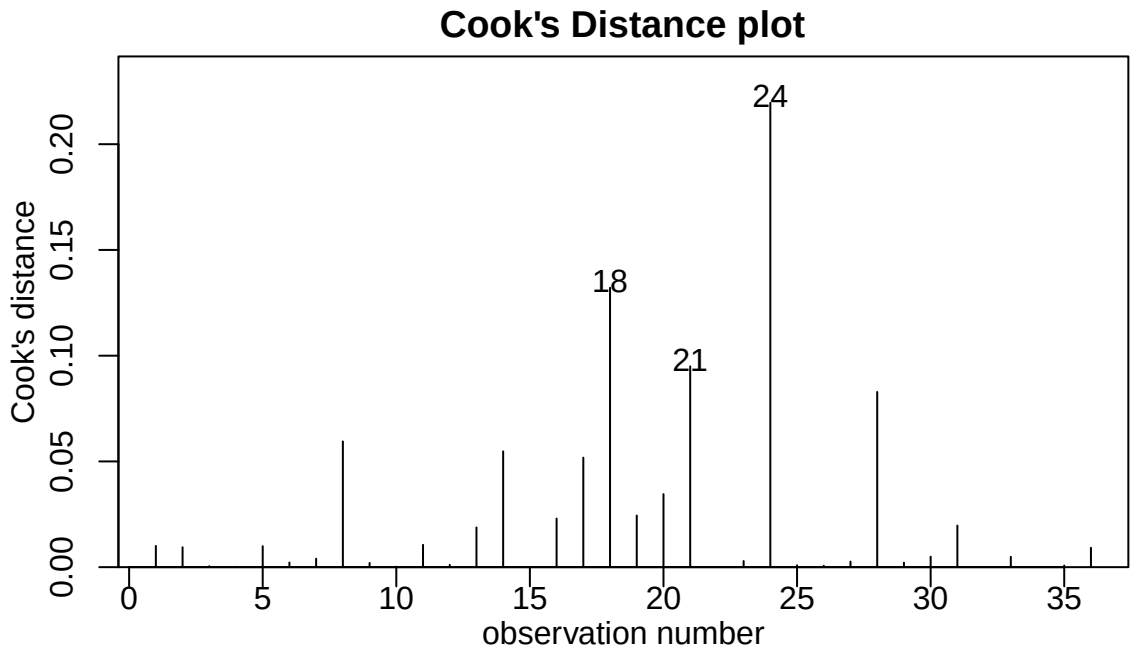
```
MSS/TSS
```

```
## [1] 0.7741266
```

Task 6

Cooks Plot of Quadratic Linear Model

```
library(s20x)
cooks20x(quad.lm)
```



Summary of Quadratic Linear Model Excluding 24th Datum

```
quad2.lm=lm(Height~BHDiameter + I(BHDiameter^2) , data=spruce.df[-24,])
summary(quad2.lm)
```

```
##
## Call:
## lm(formula = Height ~ BHDiameter + I(BHDiameter^2), data = spruce.df[-24,
##     ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.11233 -0.48227  0.01253  0.71727  2.59146
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -0.341500   2.068479  -0.165     0.87
## BHDiameter       1.564793   0.226102   6.921 7.78e-08 ***
## I(BHDiameter^2) -0.029242   0.006114  -4.782 3.74e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.266 on 32 degrees of freedom
## Multiple R-squared:  0.8159, Adjusted R-squared:  0.8044
## F-statistic: 70.91 on 2 and 32 DF,  p-value: 1.74e-12
```