Lab3-R

Clayton Glenn

January 29, 2018

Task 1

Get Working Directory

```
getwd()
```

```
## [1] "C:/Users/cglen/Documents/Stat Methods/Labs/LAB3"
```

Task 2

Read Data from the Data File and Show the First 5 Lines
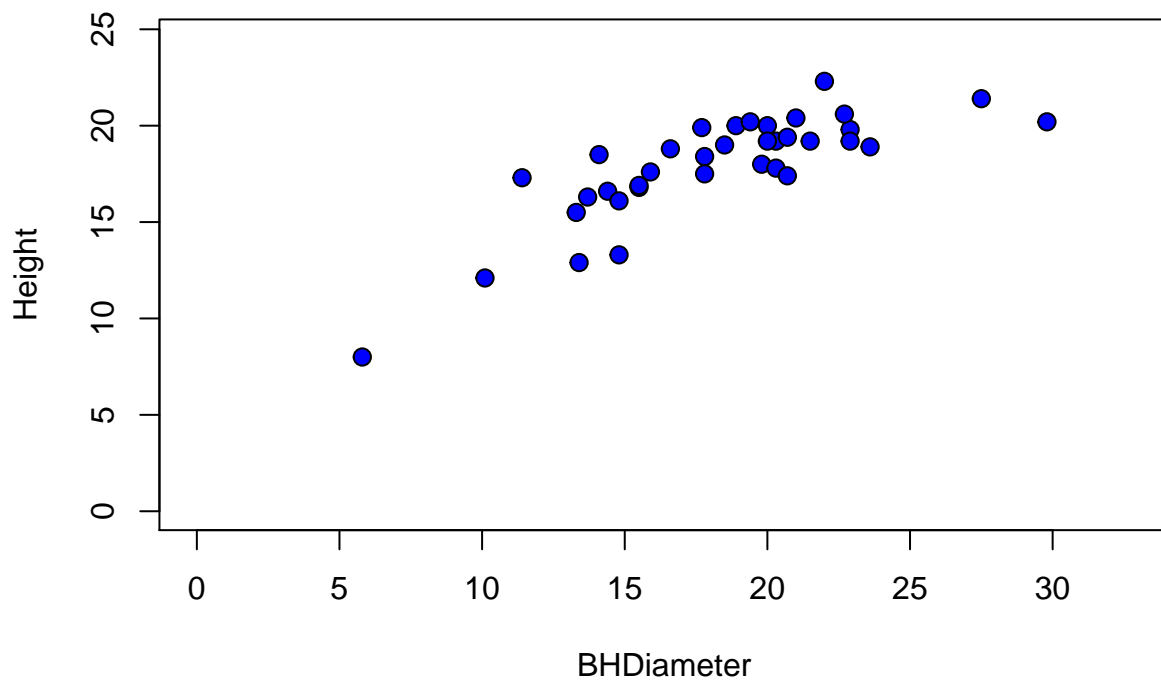
```
spruce.df <- read.table("SPRUCE.csv", header = TRUE, sep = ",")
head(spruce.df)
```

```
##   BHDiameter Height
## 1       18.9   20.0
## 2       16.6   18.8
## 3       15.5   16.8
## 4       15.5   16.9
## 5       19.4   20.2
## 6       13.7   16.3
```

Task 3

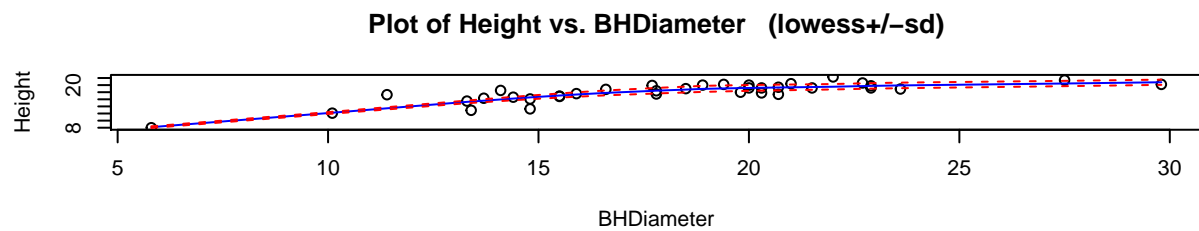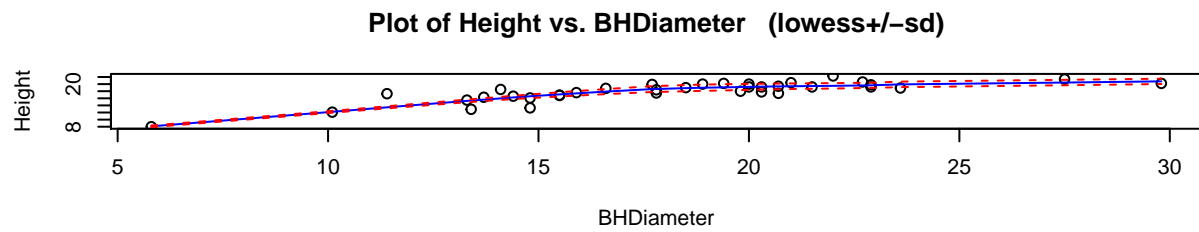### Scatter Plot of Data with X(Breast height Diameter) and Y(Height)
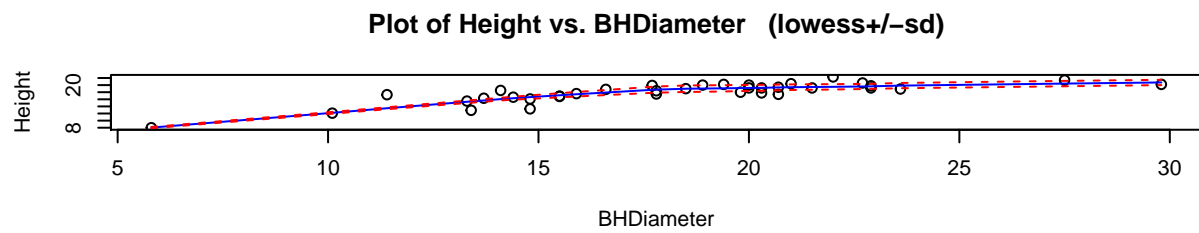
```
with(spruce.df, {
  plot(Height ~ BHDiameter, bg = "Blue", pch = 21, cex = 1.2,
  ylim = c(0, 1.1*max(Height)), xlim = c(0, 1.1*max(BHDiameter)))
} )
```

According to the Scatter Plot, a Straight Line isn't likely to fit the data.

## S20X TrendScatter Plot of Height V BHDiameter

```
library(s20x)
with(spruce.df, {
  layout(matrix(1:3, nr = 3))
  trendscatter(Height ~ BHDiameter,f=0.5)
  trendscatter(Height ~ BHDiameter,f=0.6)
  trendscatter(Height ~ BHDiameter,f=0.7)
})
```

Plot of Height vs. BHDiameter   (lowess+/−sd)

Plot of Height vs. BHDiameter   (lowess+/−sd)

Plot of Height vs. BHDiameter   (lowess+/−sd)

Scatter Plot W/ Regression Line of Height V BHDiameter

```
spruce.lm <- with(spruce.df, lm(Height~BHDiameter))
with(spruce.df,
  plot(Height ~ BHDiameter, bg = "Blue", pch = 21, ylim = c(0, 1.1*max(Height)),
  xlim = c(0, 1.1*max(BHDiameter)))
)
abline(spruce.lm)
```
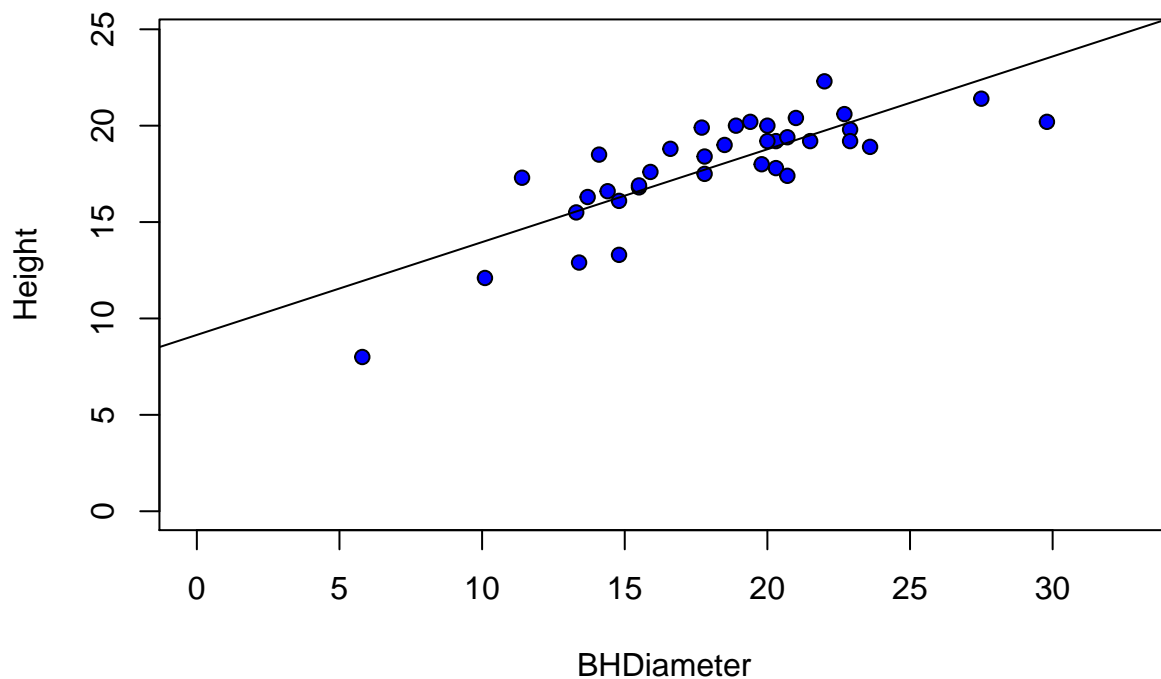
The data group follows a curve with a negative slope that tend towards zero. Therefore the data does not fit a straight line, instead, a smooth curve would fit the data better.

Task 4

Calculations of YHAT, MSS, TSS, and RSS

```
yhat <- with(spruce.df,predict(spruce.lm,data.frame(BHDiameter)))
MSS <- with(spruce.df,sum((yhat-mean(Height))^2))
TSS <- with(spruce.df,sum((Height-mean(Height))^2))
RSS <- with(spruce.df,sum((Height-yhat)^2))
```

Scatter Plots w/ Regression, Mean, and Expected

```
with(spruce.df, {
  layout(matrix(1:4,nr=2,nc=2,byrow=TRUE))

  plot(Height~BHDiameter,bg="Blue",pch=21,ylim=c(0,1.1*max(Height)),xlim=c(0,1.1*max(BHDiameter)))
  abline(spruce.lm)
  mtext("Linear Model")

  plot(Height~BHDiameter,bg="Blue",pch=21,ylim=c(0,1.1*max(Height)),xlim=c(0,1.1*max(BHDiameter)))
```

```
    segments(BHDiameter,Height,BHDiameter,yhat)
    abline(spruce.lm)
    mtext("Diff between Point and Expected")

    plot(Height~BHDiameter,bg="Blue",pch=21,ylim=c(0,1.1*max(Height)),xlim=c(0,1.1*max(BHDiameter)))
    abline(spruce.lm)
    abline(h=mean(Height))
    segments(BHDiameter,mean(Height),BHDiameter,yhat,col="Red")
    mtext("Diff between Linear Model and Mean")

    plot(Height~BHDiameter,bg="Blue",pch=21,ylim=c(0,1.1*max(Height)),xlim=c(0,1.1*max(BHDiameter)))
    abline(h=mean(Height))
    segments(BHDiameter,Height,BHDiameter,mean(Height),col="Green")
    mtext("Diff between Point and Mean")
})
```
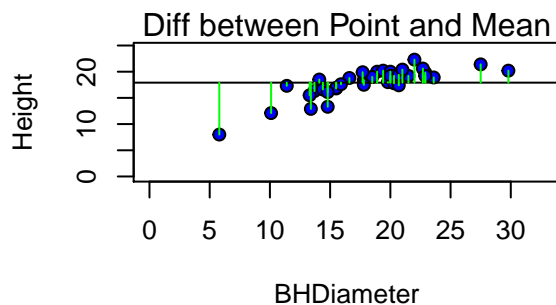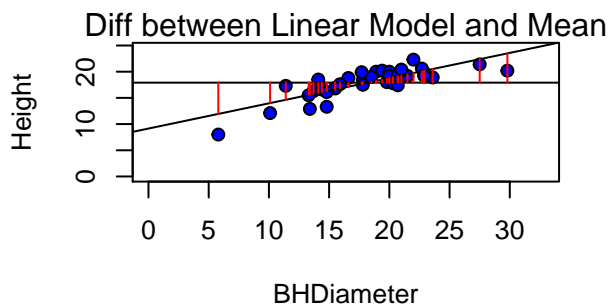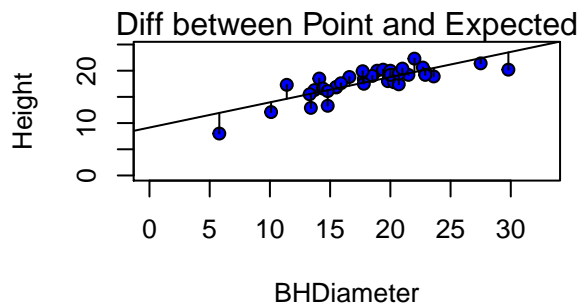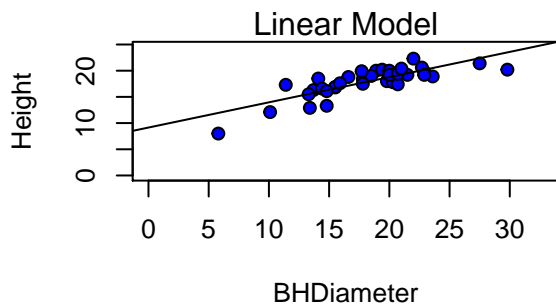

Show All Sums of Squares and Calculations

TSS

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

MSS

```
## [1] 183.2447
```

```
RSS
```

[1] 95.70281

```
MSS/TSS
```

[1] 0.6569146

```
MSS+RSS
```

[1] 278.9475

Total does in Fact Equal the Sum of Model and Residual. Total Sum of Squares / Model Sum of Squares is the slope of the Linear Model.

Task 5

Summary of Linear Model

```
summary(spruce.lm)
```

```
##
## Call:
## lm(formula = Height ~ BHDiameter)
##
## Residuals:
##     Min     1Q  Median      3Q     Max
## -3.9394 -0.9763  0.2829  0.9950  2.6644
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.14684    1.12131   8.157 1.63e-09 ***
## BHDiameter   0.48147    0.05967   8.069 2.09e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.678 on 34 degrees of freedom
## Multiple R-squared:  0.6569, Adjusted R-squared:  0.6468
## F-statistic:  65.1 on 1 and 34 DF,  p-value: 2.089e-09
```

Value of Slope is .6569. Value of Intercept is 9.14684. The equation of the Fitted Line is y=.6569x+9.14684.

Coefficients of Linear Model

```
coef(spruce.lm)
```

```
## (Intercept)  BHDiameter
##   9.1468390   0.4814743
```

Regression Analysis

```
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
```

Predict Height @ 15, 18, 20

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

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

Task 6

GGPLOT of Data, Height V BHDiameter W/ Regression

```
library(ggplot2)
  g <- ggplot(spruce.df, aes(x=BHDiameter,y=Height,colour=BHDiameter))
  g <- g+geom_point() + geom_line()+ geom_smooth(method="lm")
  g+ggtitle("Height Vs BHDiameter")
```

Height Vs BHDiameter