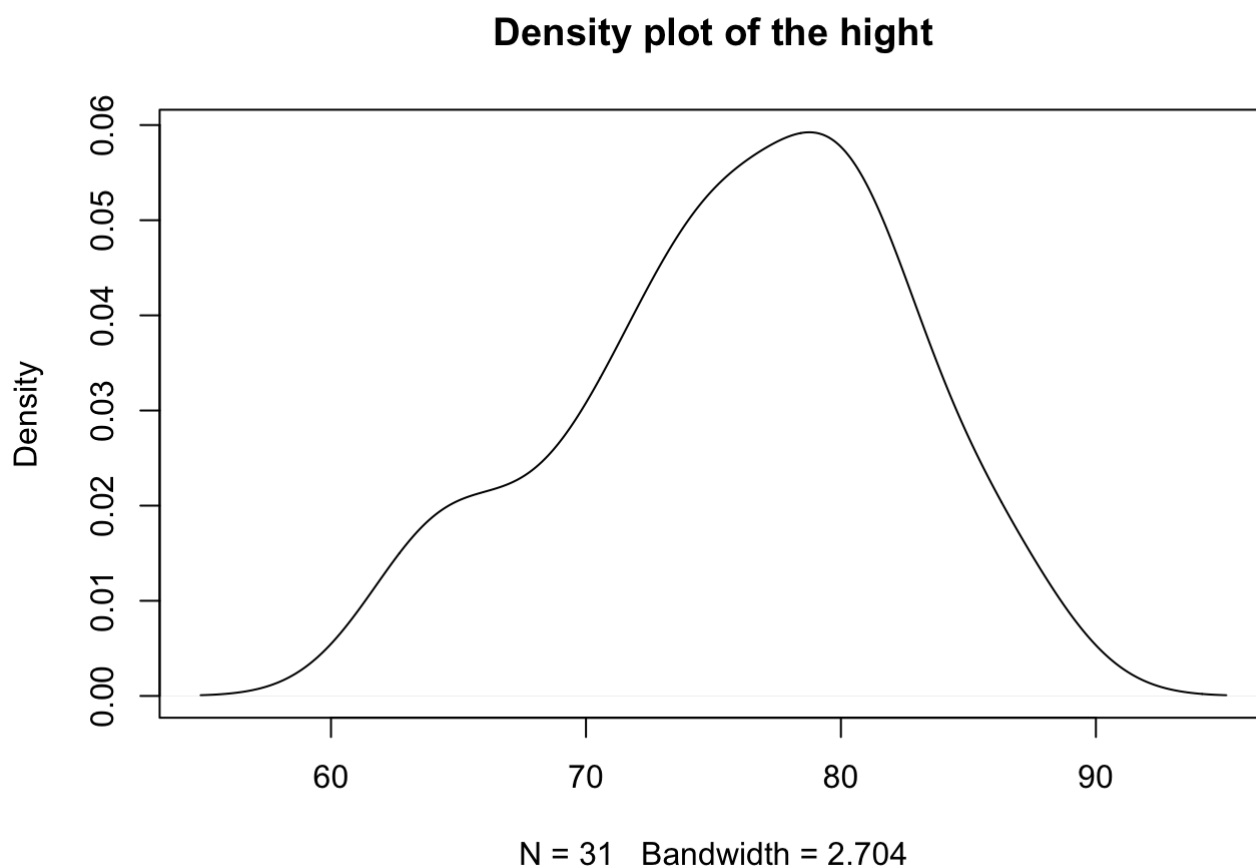# Homework1

*Yi Chen*

*1/23/2020*

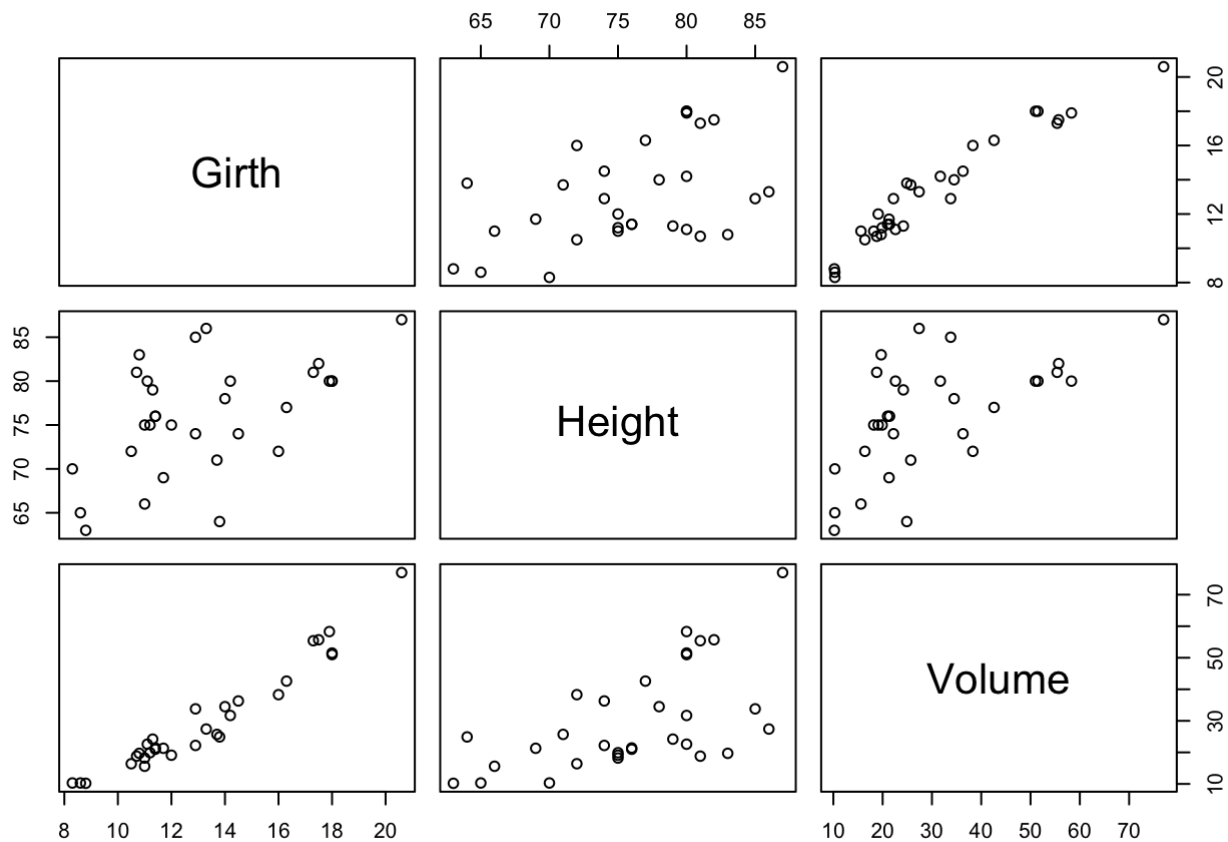## 1. Use the built-in dataset trees to do the following:

### a) Obtain the density plot of the Height variable.

```
data(trees)
attach(trees)
plot(density(trees$Height), main="Density plot of the hight")
```

**Density plot of the hight**



N = 31   Bandwidth = 2.704

### b) Obtain a scatterplot matrix of all 3 variables.

```
pairs(trees[, 1:3])
```

## c) Obtain the mean vector, the variance-covariance matrix ($S_n$), and the correlation matrix of all 3 variables.

```
# column mean
colMeans(trees)
```

```
##     Girth    Height    Volume
## 13.24839 76.00000 30.17097
```

```
# variance-covariance matrix
cov(trees)
```

```
##              Girth    Height     Volume
## Girth     9.847914 10.38333   49.88812
## Height   10.383333 40.60000   62.66000
## Volume   49.888118 62.66000  270.20280
```

```
# correlation matrix
cor(trees)
```

```
##              Girth      Height      Volume
## Girth   1.0000000 0.5192801 0.9671194
## Height 0.5192801 1.0000000 0.5982497
## Volume 0.9671194 0.5982497 1.0000000
```

# 2. Do exercise 3 on p. 38.

```
x1 <- c(9,2,6,5,8)
x2 <- c(12,8,6,4,10)
x3 <- c(3,4,0,2,1)
X <- cbind(x1,x2,x3)
```

The array of measure mean is

```
apply(X,2,mean)
```

```
## x1 x2 x3
##  6  8  2
```

The variance-covariance matrix is

```
cov(X)
```

```
##         x1    x2     x3
## x1   7.50   5.0 -1.75
## x2   5.00  10.0  1.50
## x3  -1.75   1.5  2.50
```

Notice: the covariance in this function is based on n - 1 degrees of freedom instead of n.

The correlation matrix is

```
cor(X)
```

```
##              x1          x2          x3
## x1   1.0000000 0.5773503 -0.4041452
## x2   0.5773503 1.0000000  0.3000000
## x3  -0.4041452 0.3000000  1.0000000
```