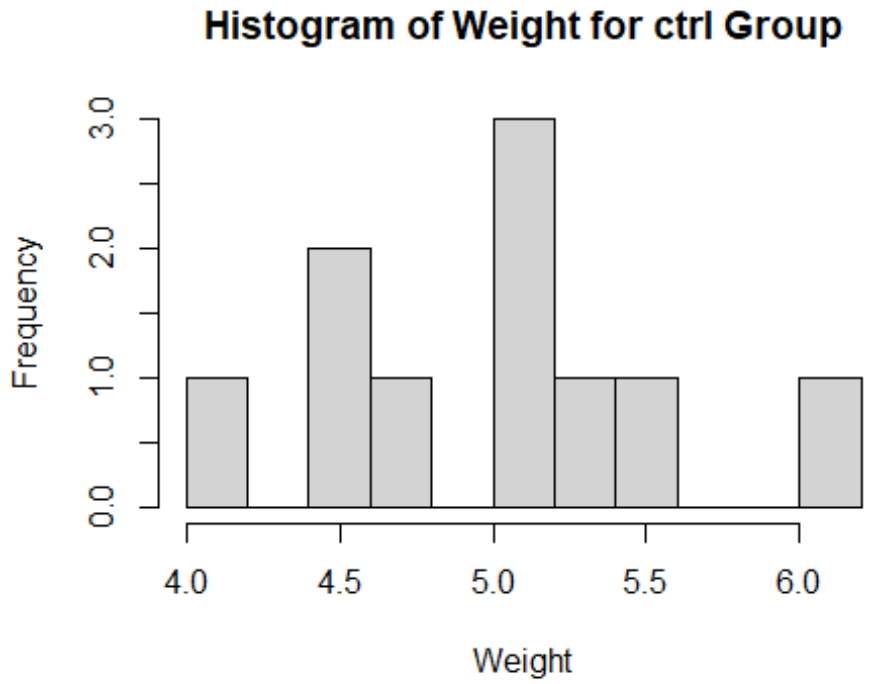
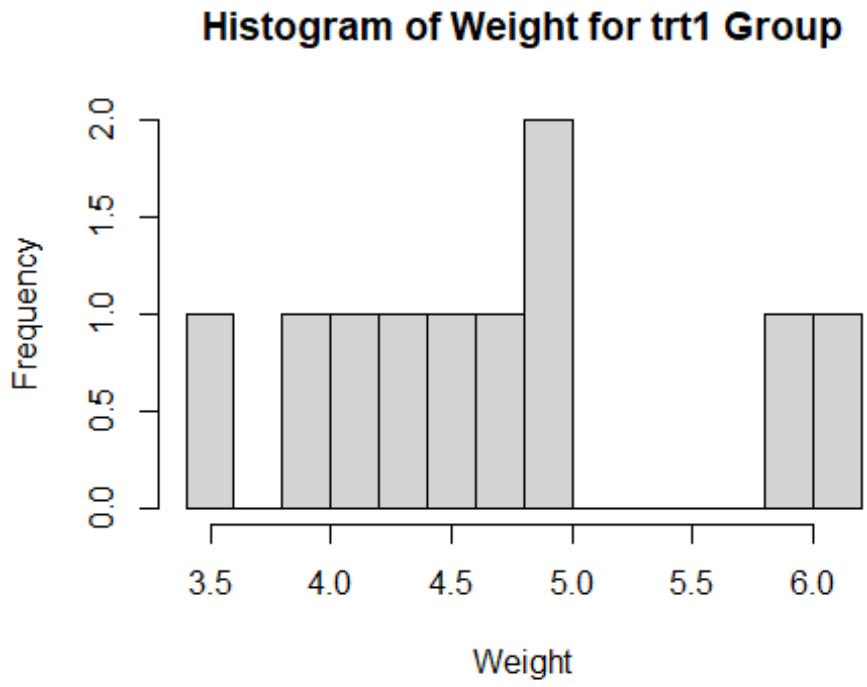
Homework 4
Syracuse University
IST 772
Summer 2021

## Question 7

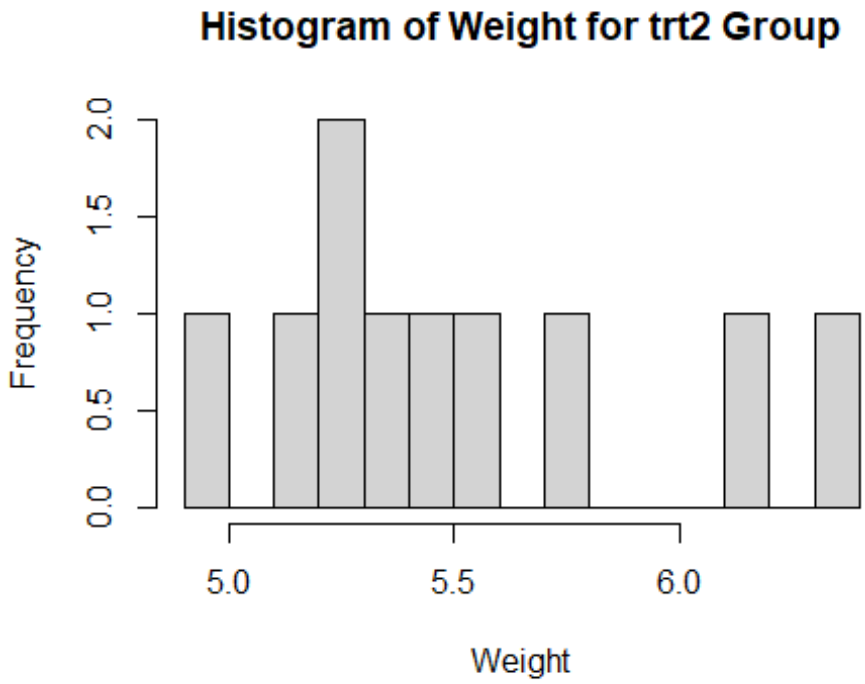
```
# summarize plant growth data and explain the output
summary(PlantGrowth)

##      weight       group
##  Min.   :3.590   ctrl:10
##  1st Qu.:4.550   trt1:10
##  Median :5.155   trt2:10
##  Mean   :5.073
##  3rd Qu.:5.530
##  Max.   :6.310

# the summary shows the min, 1st quartile, median, mean, 3rd quartile,
# and max of the weight variable of the plant growth dataset. It also
# shows that there are 3 groups each with 10 observations.

# create a histogram of the control group
hist(PlantGrowth[which(PlantGrowth$group == 'ctrl'),
                 which(colnames(PlantGrowth) == 'weight')],
                 main = 'Histogram of Weight for ctrl Group',
                 xlab = 'Weight', ylab = 'Frequency', breaks = 10)
```



```r
# create a histogram of the trt1 group
hist(PlantGrowth[which(PlantGrowth$group == 'trt1'),
                 which(colnames(PlantGrowth) == 'weight')],
                 main = 'Histogram of Weight for trt1 Group',
                 xlab = 'Weight', ylab = 'Frequency', breaks = 10)
```



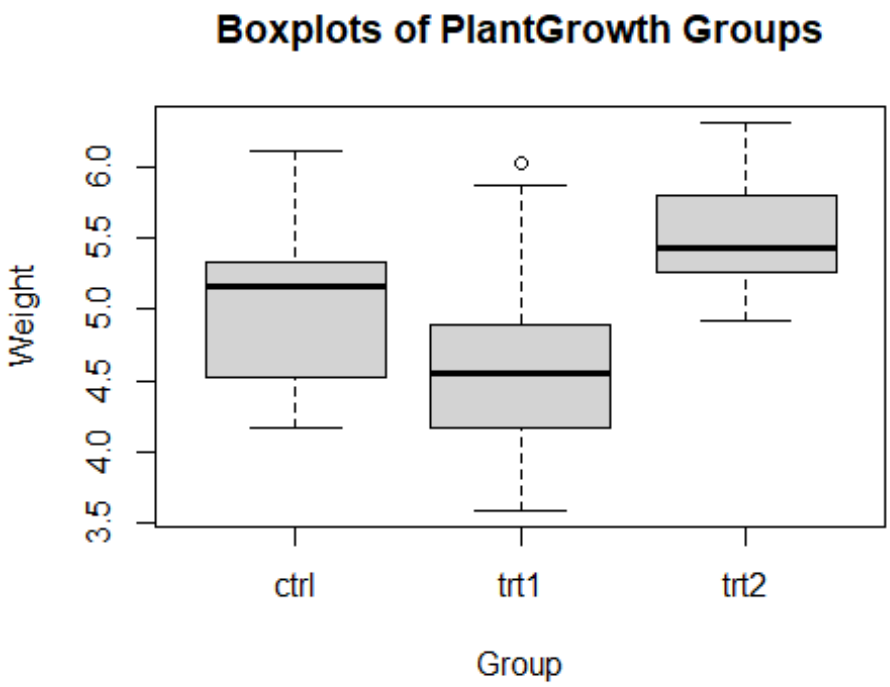

```r
# create a histogram of the trt2 group
hist(PlantGrowth[which(PlantGrowth$group == 'trt2'),
                 which(colnames(PlantGrowth) == 'weight')],
                 main = 'Histogram of Weight for trt2 Group',
                 xlab = 'Weight', ylab = 'Frequency', breaks = 10)
```




```
# the number of observations in the data set is very small, but based on
# the data that is available and by looking at the histograms, the control
# group weights are more clustered around the middle while the trt1 and
# trt2 groups are more spread out.
```

### Question 8

```
# create a box plot of the plant growth data
boxplot(formula = weight ~ group,
        data = PlantGrowth,
        main = 'Boxplots of PlantGrowth Groups',
        ylab = 'Weight',
        xlab = 'Group')
```




```
# It looks like the trt1 group tends to have lower weight than the other
groups,
# the trt2 group tends to have higher weight than the other groups, and the
# control group tends to have weight in the middle.
```

### Question 9

```
# run a t test to compare the means of ctrl and trt1 groups
t.test(PlantGrowth$weight[PlantGrowth$group == 'ctrl'],
       PlantGrowth$weight[PlantGrowth$group == 'trt1'])

##
##  Welch Two Sample t-test
##
## data:  PlantGrowth$weight[PlantGrowth$group == "ctrl"] and
PlantGrowth$weight[PlantGrowth$group == "trt1"]
## t = 1.1913, df = 16.524, p-value = 0.2504
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.2875162  1.0295162
## sample estimates:
## mean of x mean of y
##     5.032     4.661

# The confidence interval resulting from the t test is a mean difference of
# -0.29 to 1.03. This means that in 95 out of 100 trials, the population
# mean difference will fall into this confidence interval and in 5 out of 100
```

```
# trials, the population mean difference will not be in this confidence
interval.
```

### Question 10

```
# run a t test to compare the means of ctrl and trt2 groups
t.test(PlantGrowth$weight[PlantGrowth$group == 'ctrl'],
       PlantGrowth$weight[PlantGrowth$group == 'trt2'])

##
##  Welch Two Sample t-test
##
## data:  PlantGrowth$weight[PlantGrowth$group == "ctrl"] and
PlantGrowth$weight[PlantGrowth$group == "trt2"]
## t = -2.134, df = 16.786, p-value = 0.0479
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.98287213 -0.00512787
## sample estimates:
## mean of x mean of y
##     5.032     5.526

# The confidence interval resulting from the t test is a mean difference of
# -0.98 to -0.01. This means that in 95 out of 100 trials, the population
# mean difference will fall into this confidence interval and in 5 out of 100
# trials, the population mean difference will not be in this confidence
interval.
```