IST 777

Joe Hernandez

7/31/2019

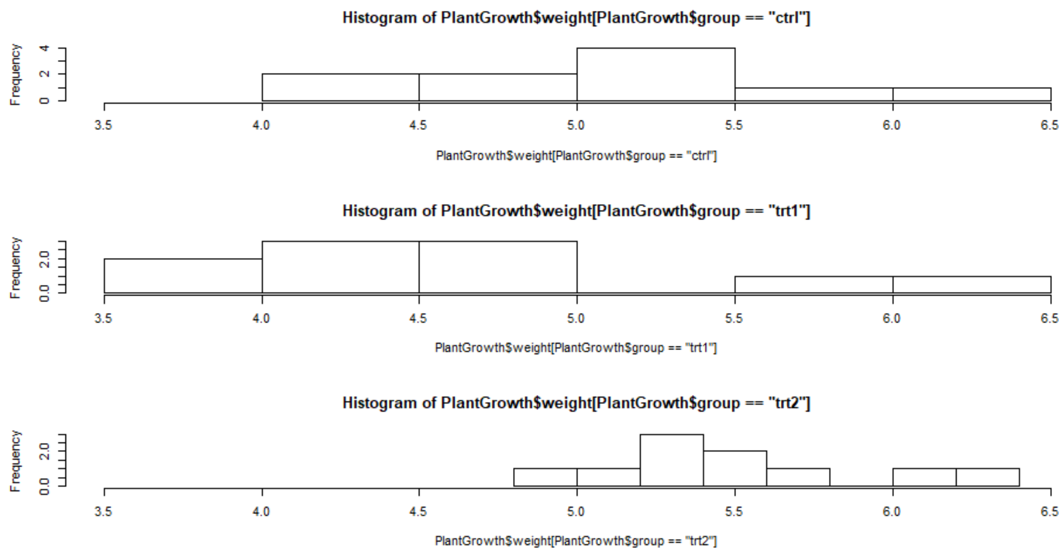
Homework #4

7. Run summary() command and explain output

```
     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 output from the summary shows two variables of weight and groups. The group variable shows 10 of each category (Control, Treatment 1, and Treatment 2), whereas, the weight variable shows the different weights for each category. The weights show the the following weight Minimum: 3.590, 1st quartile: 4.550, Median: 5.155, Mean: 5.073, 3rd quartile: 5.530, and Maximum 6.310. This information shows the data weight ranges 2.72 and the measures of central tendencies are fairly close to one another.

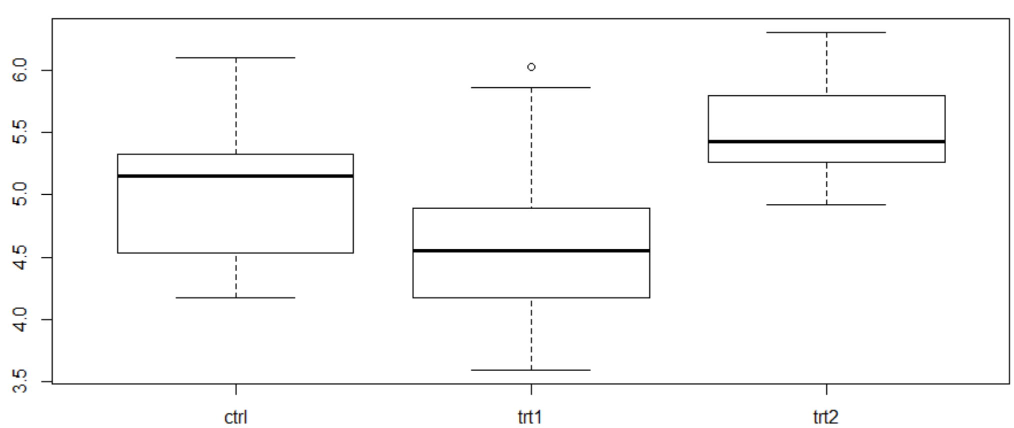
Put graph in same display and add range function for better comparison
```
par(mfrow=c(3,1))
```
Create Histogram of control group
```
hist(PlantGrowth$weight[PlantGrowth$group == "ctrl"],
xlim=c(3.5,6.5))
```
Create Histogram of Treatment 1 group
```
hist(PlantGrowth$weight[PlantGrowth$group == "trt1"],
xlim=c(3.5,6.5))
```
Create Histogram of Treatment 2 group
```
hist(PlantGrowth$weight[PlantGrowth$group == "trt2"],
xlim=c(3.5,6.5))
```
Reset graph display back to normal
```
par(mfrow=c(1,1))
```



The histogram shows very little information of the raw data. Given it is 10 records for each group the histogram under the control has what appears as a normal to uniform distribution of the data. Treatment 1 looks close to Poisson to uniform in its distribution and Treatment 2 has a more Poisson distribution.

8. Create boxplot of plant growth data

```
boxplot(weight ~ group, data = PlantGrowth)
```



Looking at the boxplots of the three groups we establish that the control group is the standard with a range weight > 4 to < 6.5 (~2) and a median weight just above 5. Under Treatment 1 we see there is higher variability in the range with the minimum weight above 3.5 and maximum around 6 (~2.5) and the median is lower than the control at a weight around 4.5. Treatment 2 shows an increase in weight when compared to the control group with a tighter range (~5 minimum - <6.5 maximum) and a slightly higher median weight just below 5.5. Lastly, when comparing the three groups each successive treatment showed a tighter interquartile range as well with treatment 2 as the tightest.

9. Run t test compare means of Control and Treatment 1

```
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
```

We constructed a 95% confidence interval around the mean difference between the control and Treatment 1 group, which ranged weight from -0.2875162 to 1.0295162. Note that this confidence interval may or may not contain the true population value.

10. Run t test compare means of Control and Treatment 2

```
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
```

We constructed a 95% confidence interval around the mean difference between the control and Treatment 1 group, which ranged weight from -0.98287213 to -0.00512787. Note that this confidence interval may or may not contain the true population value.