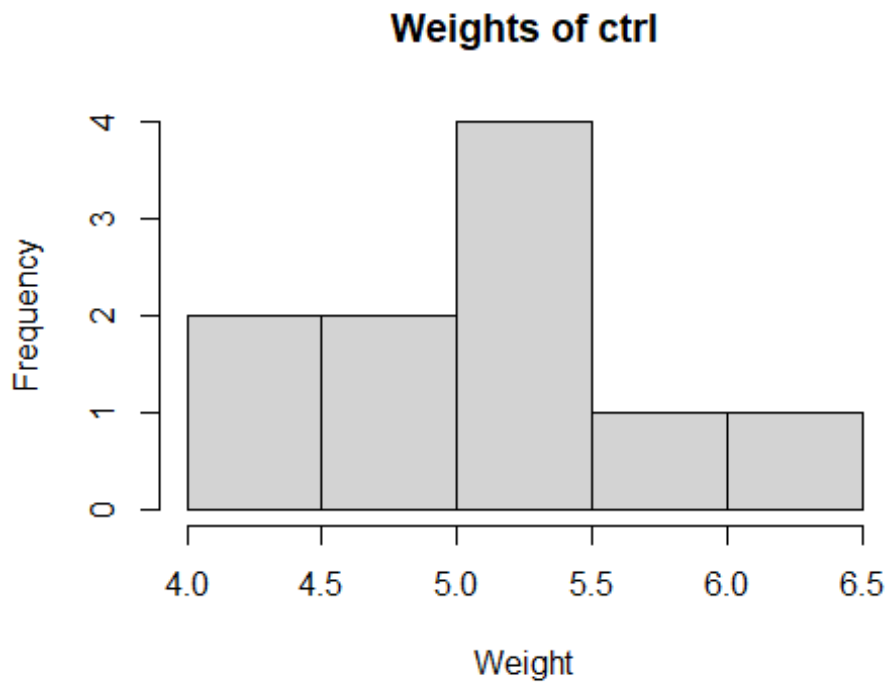
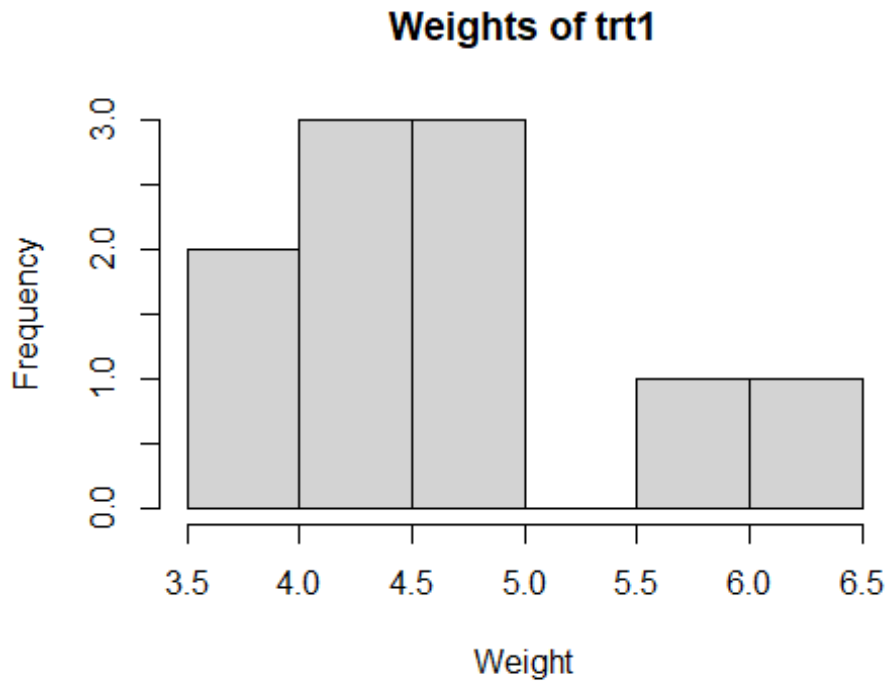
HW4_Videtti

```
##7. The built-in PlantGrowth data set contains three different groups, each representing a different plant food diet (you may need to type data(PlantGrowth) to activate it). The group labeled “ctrl” is the control group, while the other two groups are each a different type of experimental treatment. Run the summary() command on PlantGrowth and explain the output. Create a histogram of the ctrl group. As a hint about R syntax, here is one way that you can access the ctrl group data: PlantGrowth$weight[PlantGrowth$group==”ctrl”] Also create histograms of the trt1 and trt2 groups. What can you say about the differences in the groups by looking at the histograms?

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

#This shows summary statistics for each of our two columns, weight and group. We see that weight is a numerical variable because we see min, max, median, mean, 1st quartile, and 3rd quartile. group seems to be a categorical variable, which is intuitive given the description of the PlantGrowth data set in the question. We see that group has 3 values, ctrl, trt1, and trt2, each of which show up 10 times in the data.

hist(PlantGrowth$weight[PlantGrowth$group == 'ctrl'],main = 'Weights of ctrl',xlab = 'Weight')
```



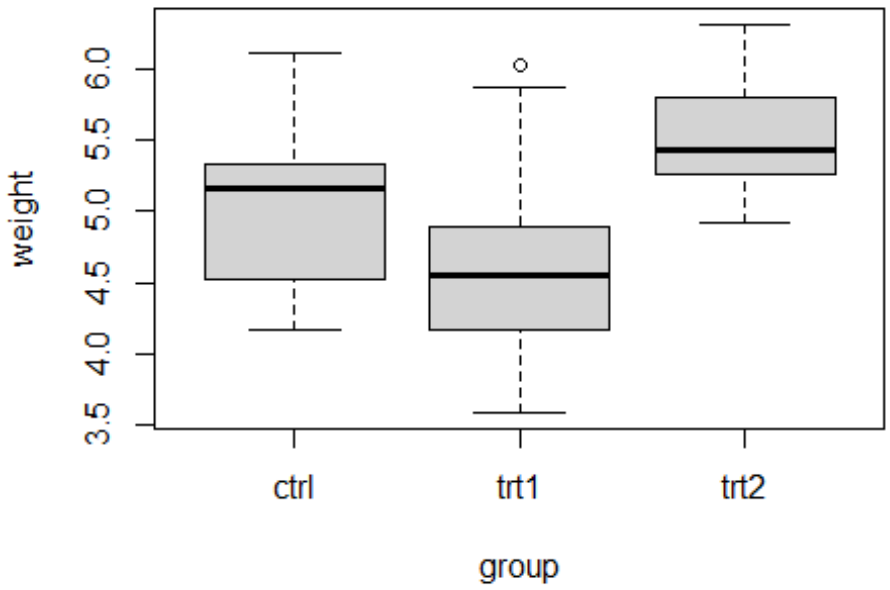

```
hist(PlantGrowth$weight[PlantGrowth$group == 'trt1'],main = 'Weights of
trt1',xlab = 'Weight')
```



```
hist(PlantGrowth$weight[PlantGrowth$group == 'trt2'],main = 'Weights of
trt2',xlab = 'Weight')
```


```
#At first glance, each of these histograms seem to indicate a mean weight in
the general vicinity 5.0. However, it appears that the mean weight is higher
for that of trt2 than in trt1, and the trt1 group looks much closer to the
ctrl group than the trt2 group. In short, weights for trt1 appear to be
slightly smaller than ctrl, and weights for trt2 appear to be larger than
ctrl, while trt1 weights appear to be closer to ctrl weights than trt2
weights are to ctrl weights.
```

```
##8. Create a boxplot of the plant growth data, using the model "weight ~
group." What can you say about the differences in the groups by looking at
the boxplots for the different groups?

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




```
#We see that our conclusions from Exercise 7 are confirmed, but we also can
now say even more. The trt1 weights cover a wider range than ctrl, while the
trt2 weights cover a smaller range than ctrl. We also see that the medians
follow a similar pattern that we hypothesized about the means in Exercise 7.
That is, the median of trt1 is lower than the median of ctrl, the median of
trt2 is higher than the median of ctrl, and the median of all three are
somewhat close to 5.0.
```

```
##9. Run a t-test to compare the means of ctrl and trt1 in the PlantGrowth
data. Report and interpret the confidence interval. Make sure to include a
carefully worded statement about what the confidence interval implies with
respect to the population mean difference between the 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
```

#Given we have a p-value of 0.2504, we do not have significant evidence to prove that the mean of the weights in the trt1 group is different than that of the ctrl group (or the mean difference in weights between the plants not given a treatment and the plants given treatment 1).

#That being said, assuming that we did have a significant t-test, the 95% confidence interval is -0.2875162 to 1.0295162. This would mean that in 95% of tests conducted, the true difference in means of weights in the control group and weights for plants given treatment 1 would be between -0.2875162 and 1.0295162.

##10. Run a t-test to compare the means of ctrl and trt2 in the PlantGrowth data. Report and interpret the confidence interval.

```
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 95% confidence interval is -0.98287213 to -0.00512787. This would mean that in 95% of tests conducted, the true difference in means of weights in the control group and weights for plants given treatment 2 would be between -0.98287213 to -0.00512787.