Assignment 2

a. Make a histogram of the variable Sepal.Width.

hist(iris$Sepal.Width)

b. Based on the histogram from #1a, which would you expect to be higher, the mean or the median? Why?

Answer: mean > median, if the graph screw to left mean will be higher than median. More outlier on the right(big value) will have make bigger average.

c. Confirm your answer to #1b by actually finding these values.

Answer: mean = 3.057333 / median = 3

mean(iris$Sepal.Width)median(iris$Sepal.Width)

d. Only 27% of the flowers have a Sepal.Width higher than _____ cm.

Answer: 3.3 cm

quarts=quantile(iris$Sepal.Width,c(.73))

e. Make scatterplots of each pair of the numerical variables in iris (There should be 6 pairs/plots).

pairs(iris[,c(1,2:4)],pch=16)

f. Based on #1e, which two variables appear to have the strongest relationship? And which two appear to have the weakest relationship?

Answer: the more vertical or horizontal or spread out, it have less realtionship between the variables.

Strogest: Petal Length vs Petal Width

Weakest: Sepal Length vs Sepal Width

---

a. Make a histogram of the variable weight with breakpoints (bin edges) at every 0.3 units, starting at 3.3.

hist(PlantGrowth$weight,breaks=seq(3.3,6.6,0.3))

b. Make boxplots of weight separated by group in a single graph.

boxplot(PlantGrowth*weight* *PlantGrowth*group)

c. Based on the boxplots in #2b, approximately what percentage of the "trt1" weights are below the minimum "trt2" weight?

Answer: 75%, trt2 min bar is just above the Q3 which is 75% of the data.

d. Find the exact percentage of the "trt1" weights that are below the minimum "trt2" weight.

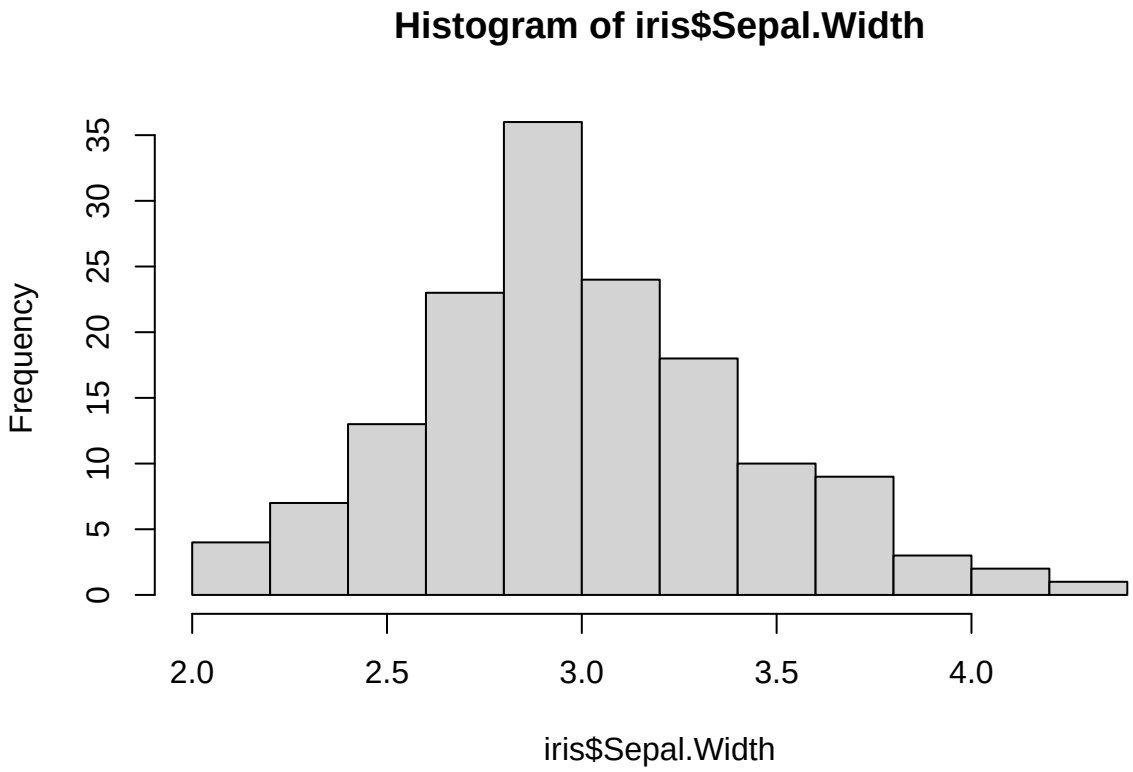
Answer: 80%, 10 out of 8 are lower than minimum(4.92) of trt2

trt1 = PlantGrowth*weight*[*PlantGrowth*group == "trt1"] trt2 = PlantGrowth*weight*[*PlantGrowth*group == "trt2"] trt1<min(trt2) sum(trt1<min(trt2))/length(trt1<min(trt2))*100

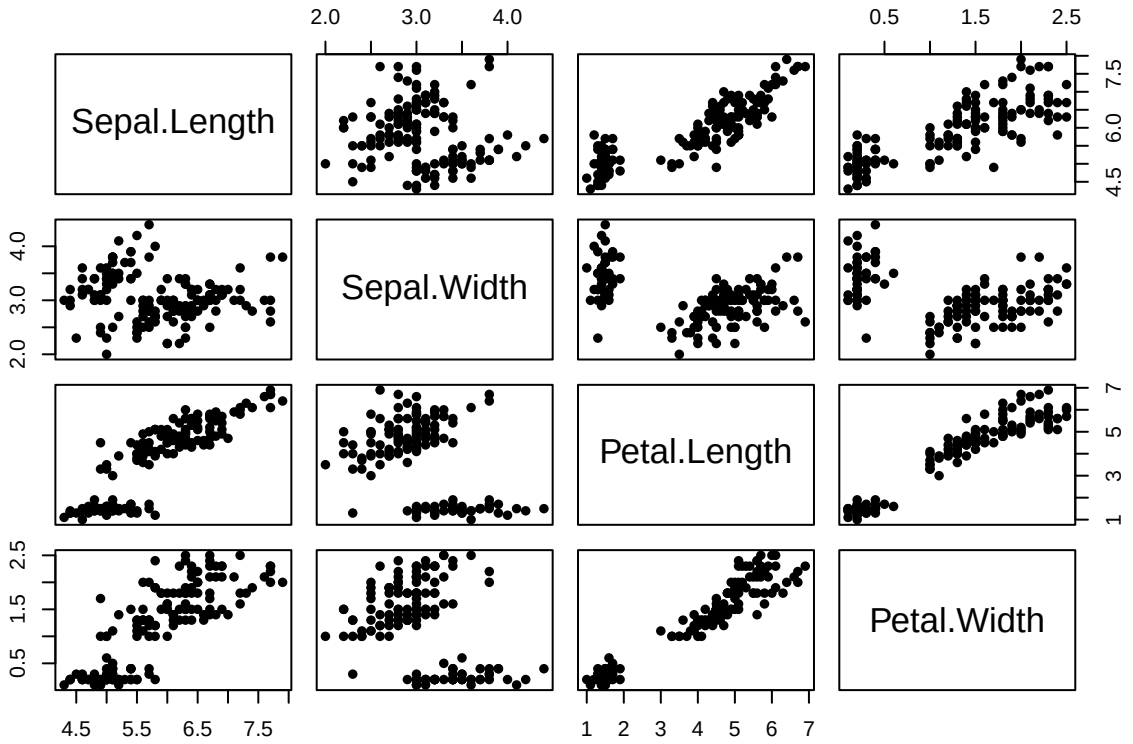
e. Only including plants with a weight above 5.5, make a barplot of the variable group. Make the barplot colorful using some color palette

Data = PlantGrowth*group*[*PlantGrowth*weight < 5.5] x = table(Data) barplot(x,col=heat.colors(3), main = "Plant Growth of weight above 5.5", xlab = 'Plant Group')

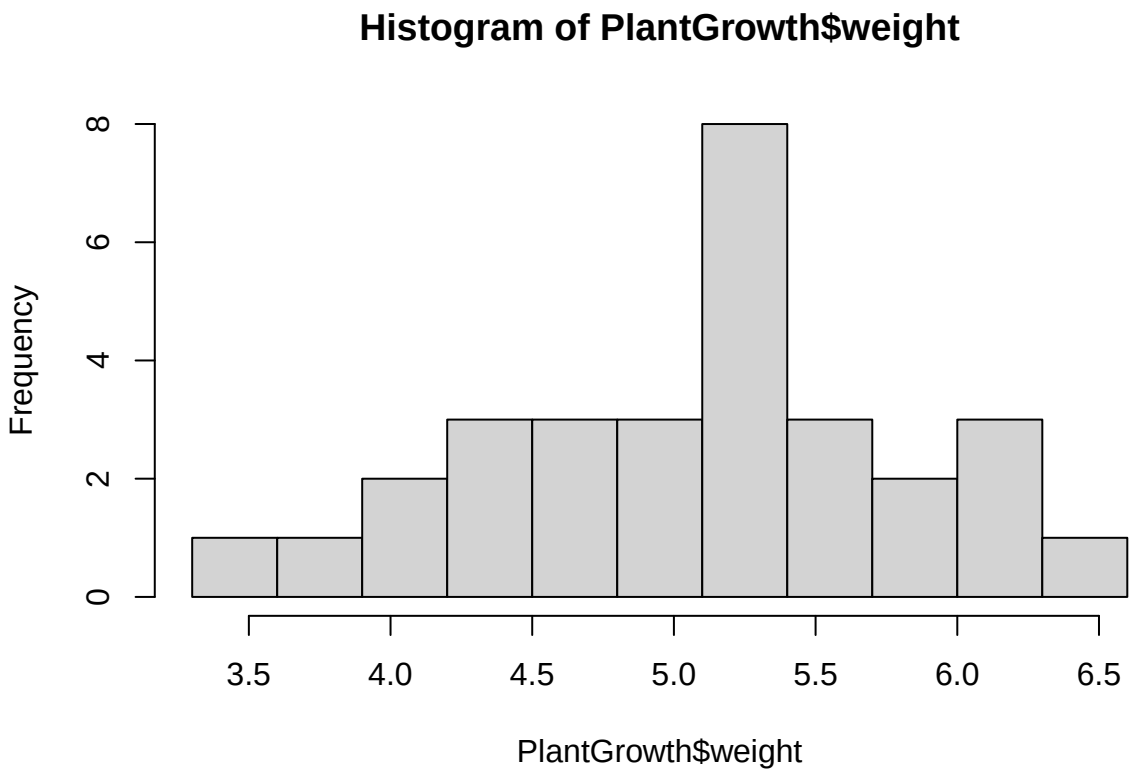
```
hist(iris$Sepal.Width)
```



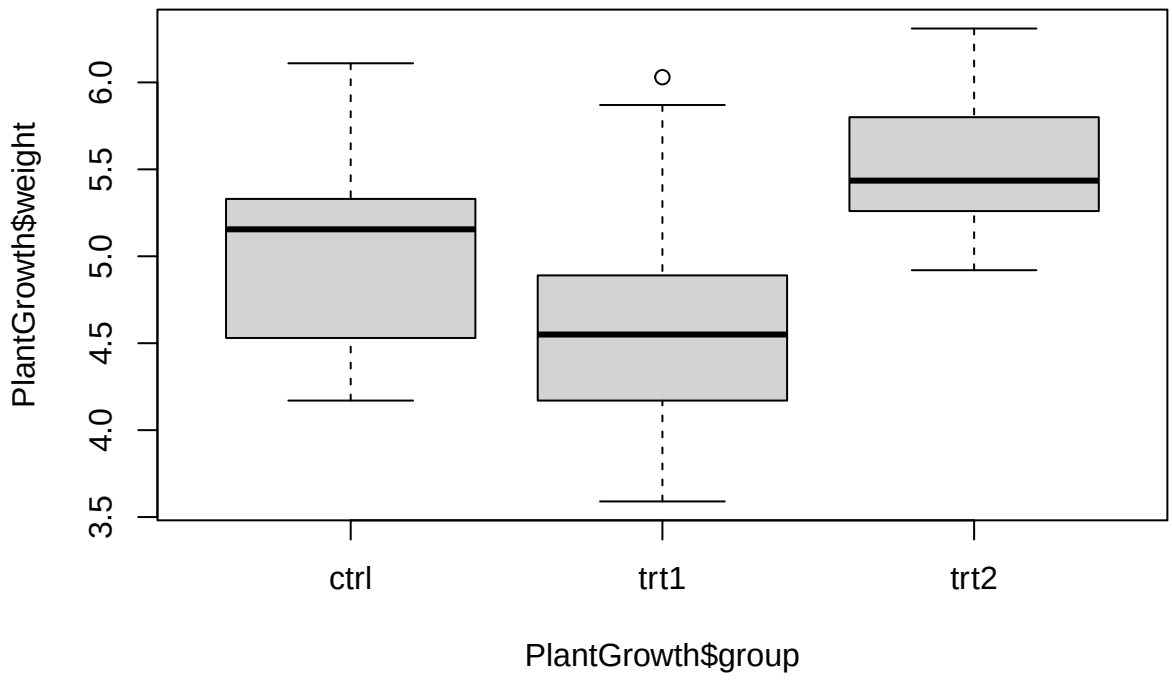

```
pairs(iris[,c(1,2:4)],pch=16)
```



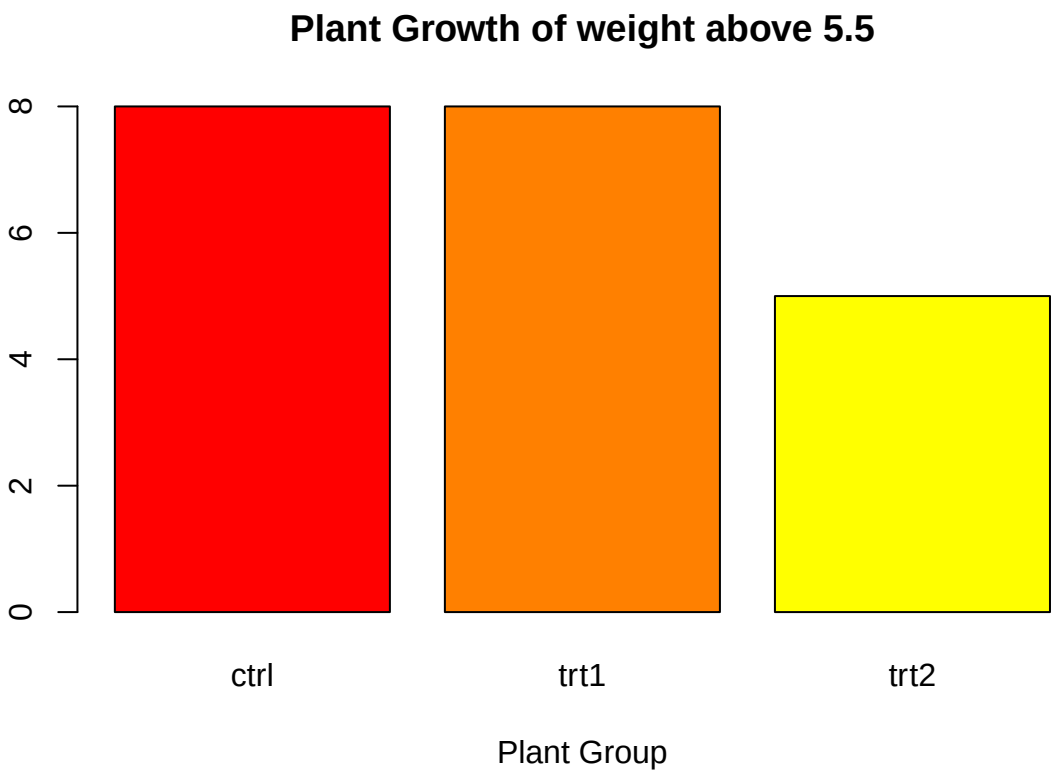

```
hist(PlantGrowth$weight,breaks=seq(3.3,6.6,0.3))
```




```
boxplot(PlantGrowth$weight~PlantGrowth$group)
```




```
barplot(table(PlantGrowth$group[PlantGrowth$weight < 5.5]),col=heat.colors(3), main = "Plant Growth of
```


Plant Growth of weight above 5.5
0
2
4
6
8
ctrl
trt1
trt2
Plant Group