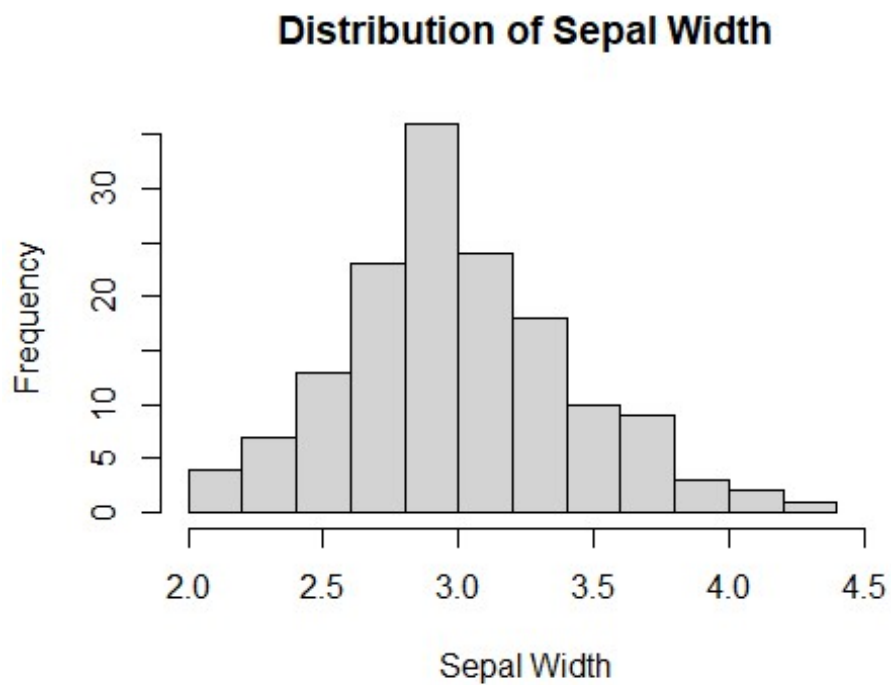


Question 1

a

```
hist(iris$Sepal.Width,  
     main = "Distribution of Sepal Width",  
     xlab = "Sepal Width",  
     xlim = c(2,4.5)  
)
```



b

Expect the mean to be higher because the distribution is slightly right-skewed

c

```
print(paste("mean:", mean(iris$Sepal.Width)))  
## [1] "mean: 3.05733333333333"  
print(paste("median:", median(iris$Sepal.Width)))  
## [1] "median: 3"
```

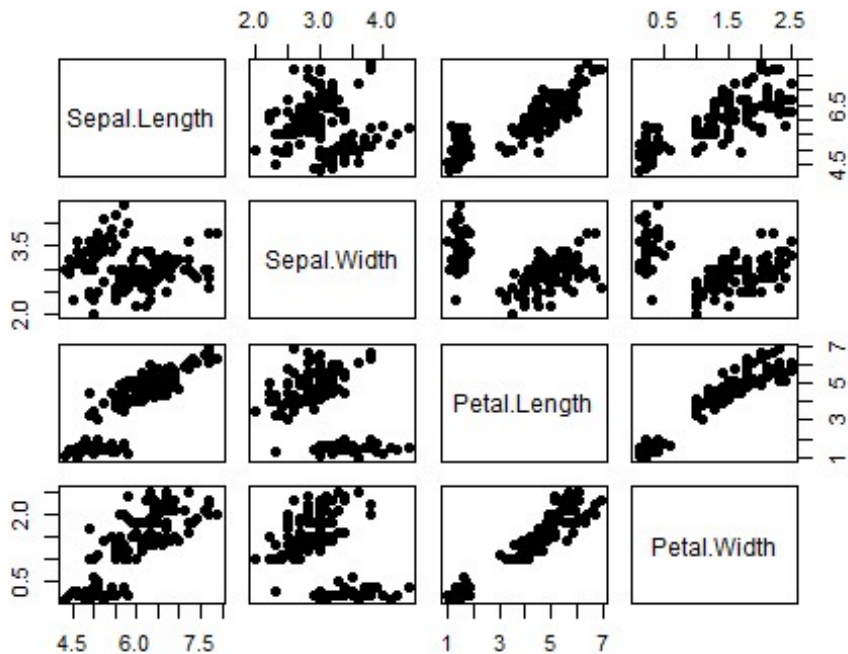
d

```
quantile(iris$Sepal.Width, .73)  
## 73%  
## 3.3
```

27% of flowers have sepal width higher than the 73rd percentile

e

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



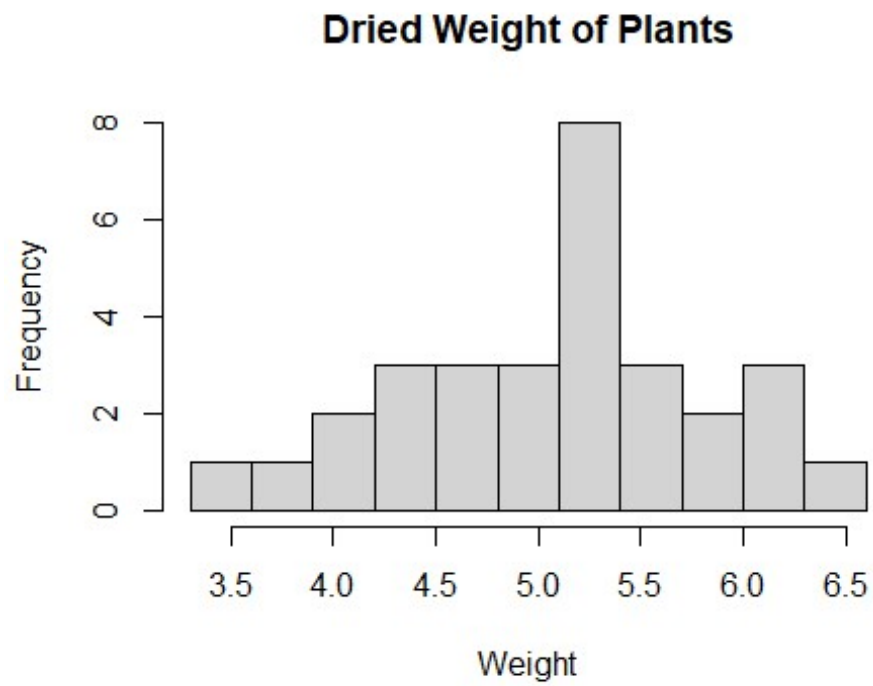
f

Petal Length and Petal Width appear to have the strongest relationship. Sepal Width and Petal Length appear to have the weakest overall relationship

Question 2

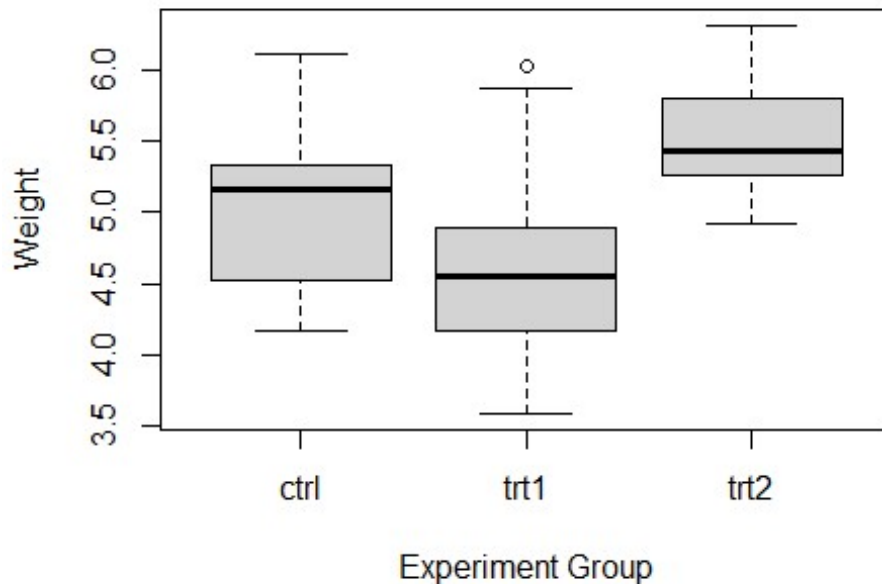
a

```
hist(PlantGrowth$weight,  
     breaks = seq(3.3, max(PlantGrowth$weight)+ 0.3, 0.3), # start at 3.3,  
     end at max+0.3, break by 0.3  
     main = "Dried Weight of Plants",  
     xlab = "Weight",  
     xlim = c(3.3,6.6)  
     )
```



b

```
boxplot(PlantGrowth$weight~PlantGrowth$group,  
        xlab = "Experiment Group",  
        ylab = "Weight"  
)
```



c

Approximately 75% of “trt1” weights are below the minimum “trt2” weight based on the boxplots

d

```
trt1_weights <- PlantGrowth$weight[PlantGrowth$group=='trt1'] # all weights
in group "trt1"
min_trt2 <- PlantGrowth$weight[PlantGrowth$group=='trt2'] # minimum weight in
group "trt2"
# counting the number of "trt1" weights below the min "trt2" weight and
dividing by n to find percentage
(sum(trt1_weights < min_trt2)) / length(trt1_weights)

## [1] 0.8
```

e

```
barplot(table(PlantGrowth$group[PlantGrowth$weight > 5.5]),
  main = "Number of Plants Heavier than 5.5 lbs",
  xlab = "Experiment Group",
  ylab = "Number of Plants",
  col = hcl.colors(3, palette = "Dynamic")
)
```

