Week 3 Assignment – Chase Werfel

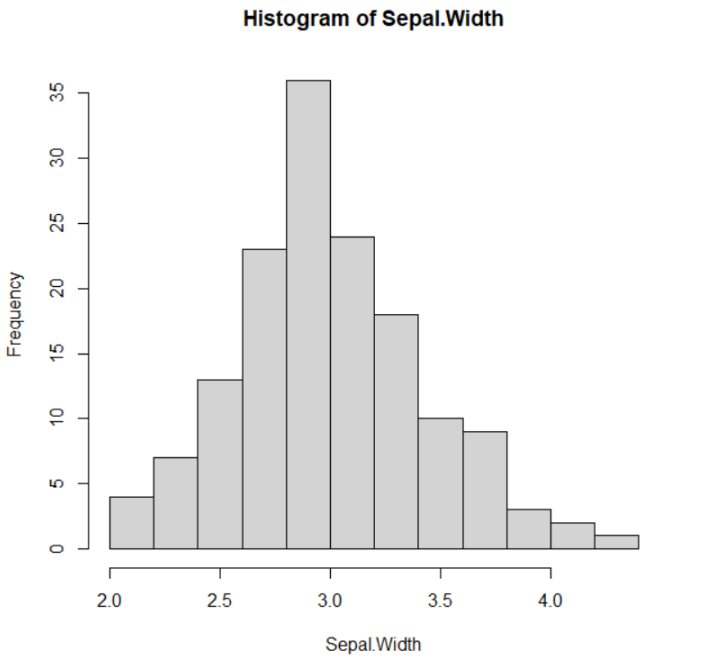
Question 1: Using the iris dataset…

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

Answer:

>attach(iris)

>hist(Sepal.Width)



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

Answer: I would expect the mean to be higher than the median because the histogram

is slightly right-skewed

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

Answer: Mean = 3.057333, Median = 3

>mean(Sepal.Width)

[1] 3.057333

>median(Sepal.Width)

[1] 3

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

Answer: 2.8 cm

> quantile(Sepal.Width,c(0.27))

27%

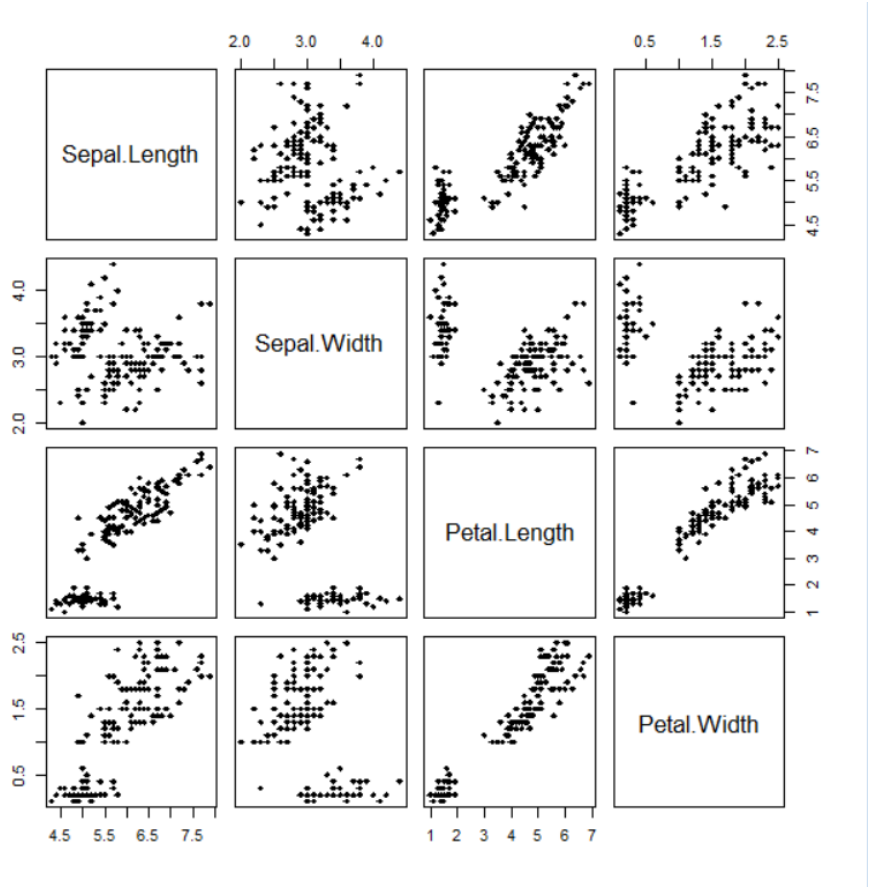
2.8

e. Make scatterplots of each pair of the numerical variables in iris

(There should be 6 pairs/plots).

Answer:

>pairs(iris[,c(1:4)],pch=18)



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

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

Question 2: Using the PlantGrowth dataset…

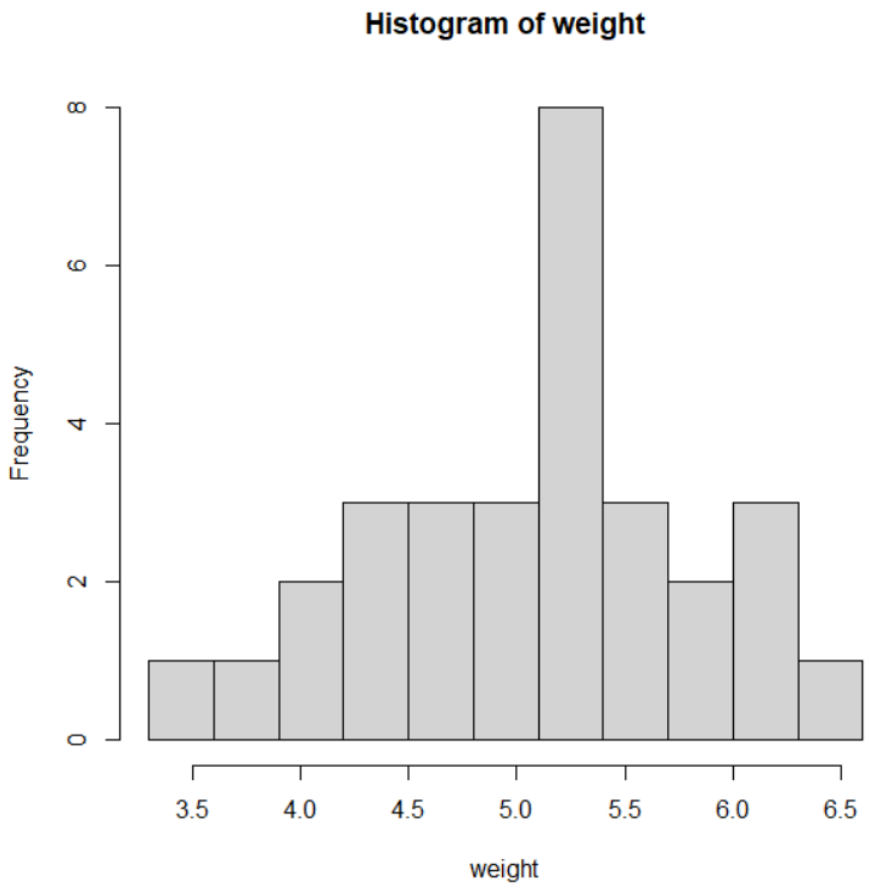
a. Make a histogram of the variable weight with breakpoints (bin edges) at every 0.3 units, starting at 3.3.

Answer:

>attach(PlantGrowth)

>binedges <- seq(from = 3.3, to = max(weight) + 0.3, by = 0.3)

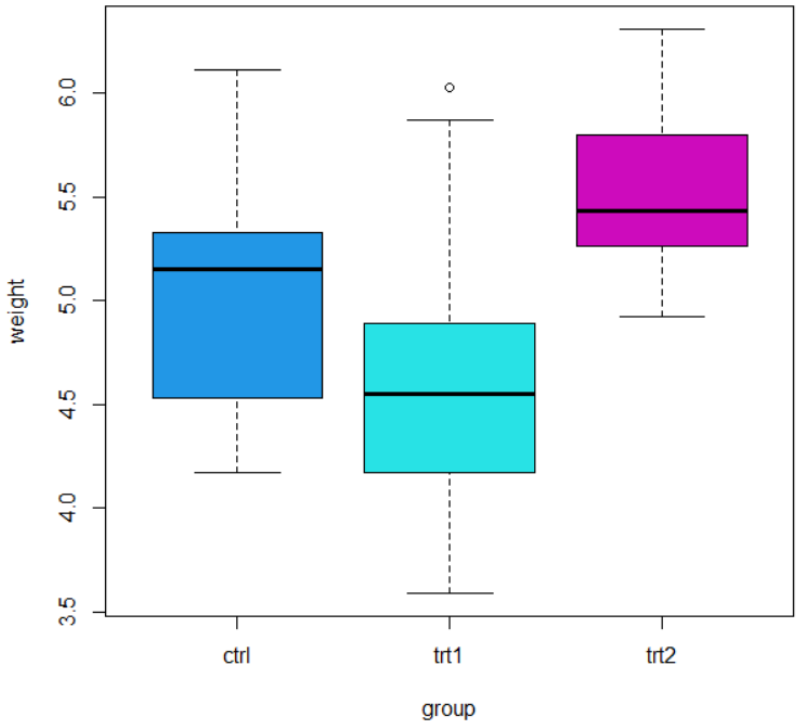
>hist(weight, breaks = binedges)



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

Answer:

>boxplot(weight~group, col=4:6)



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

Answer: 75%. The minimum trt2 weight is very close to the Q3 boundary of trt1.

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

Answer: 80%

>sum(group == "trt1" & weight < min(weight[group == "trt2"]))/sum(group == "trt1")

[1] 0.8

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.

Answer:

```
> PlantGrowth[weight>5.5, ]
   weight  group
2    5.58    ctrl
4    6.11    ctrl
15   5.87    trt1
17   6.03    trt1
21   6.31    trt2
23   5.54    trt2
28   6.15    trt2
29   5.80    trt2
```

```
>barplot(table(PlantGrowth[weight>5.5, ]), col = 4:6, ylab = "Number of Plants w/ Weight > 5.5", xlab = "Group")
```