

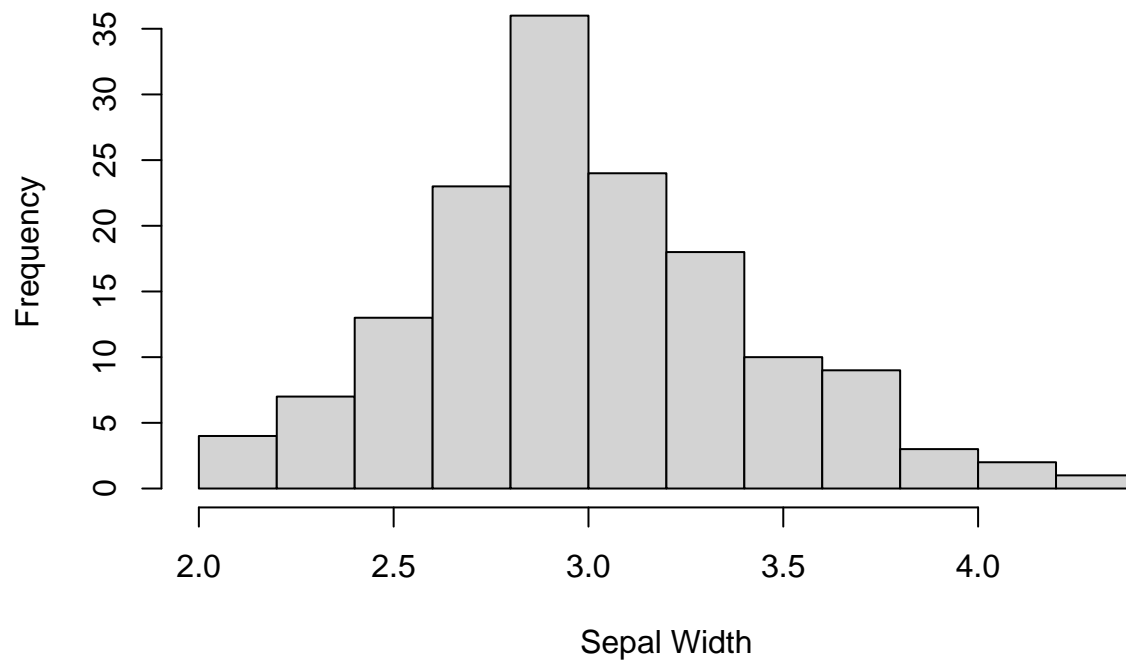
Week 3 Assignment

Emily Kolesar

2025-09-03

Part 1

```
hist(iris$Sepal.Width,  
     main=element_blank(),  
     xlab="Sepal Width")
```



b) I would expect the mean to be ever so slightly higher than the median because it is very slightly right skewed

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

```
## [1] 3.057333
```

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

```
## [1] 3
```

c) The mean is slightly greater than the median

```
100-27
```

```
## [1] 73
```

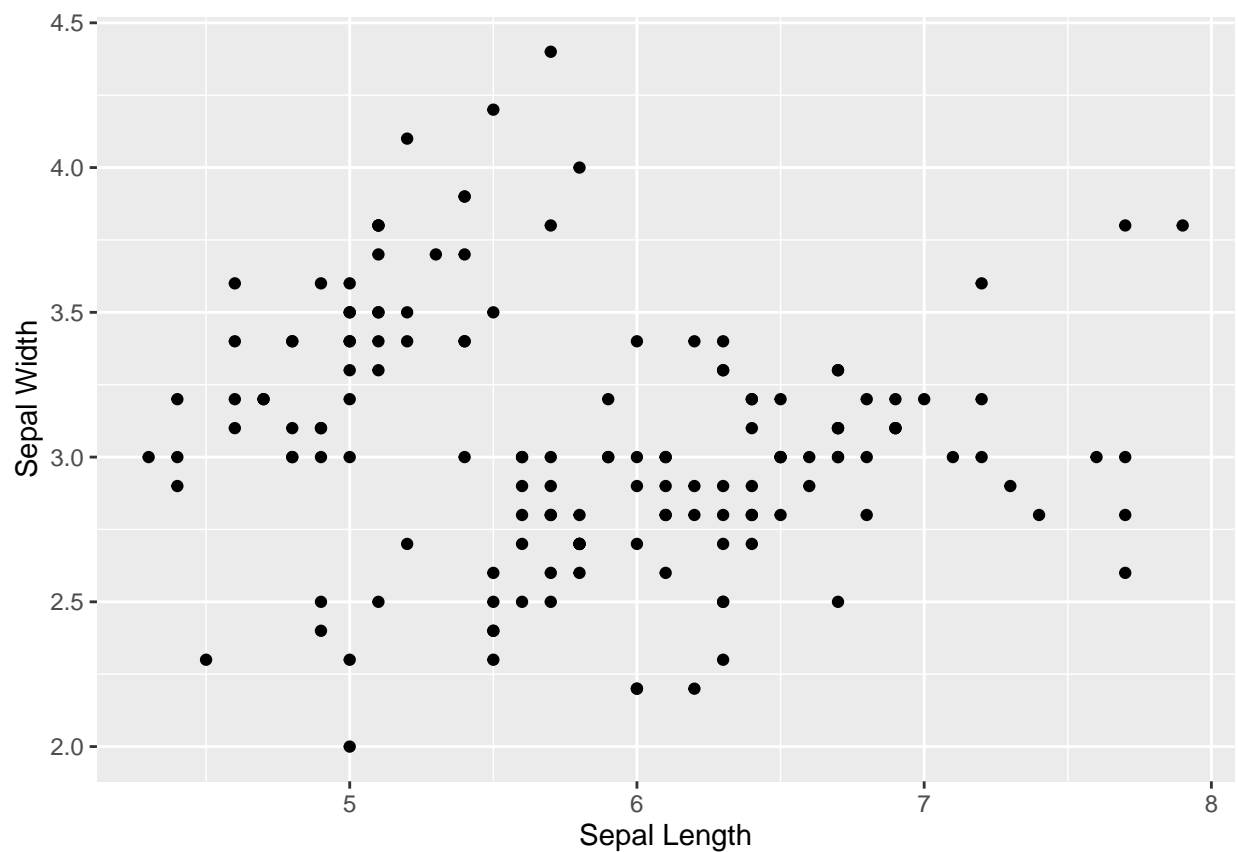
```
quantile(iris$Sepal.Width, 0.73)
```

```
## 73%
```

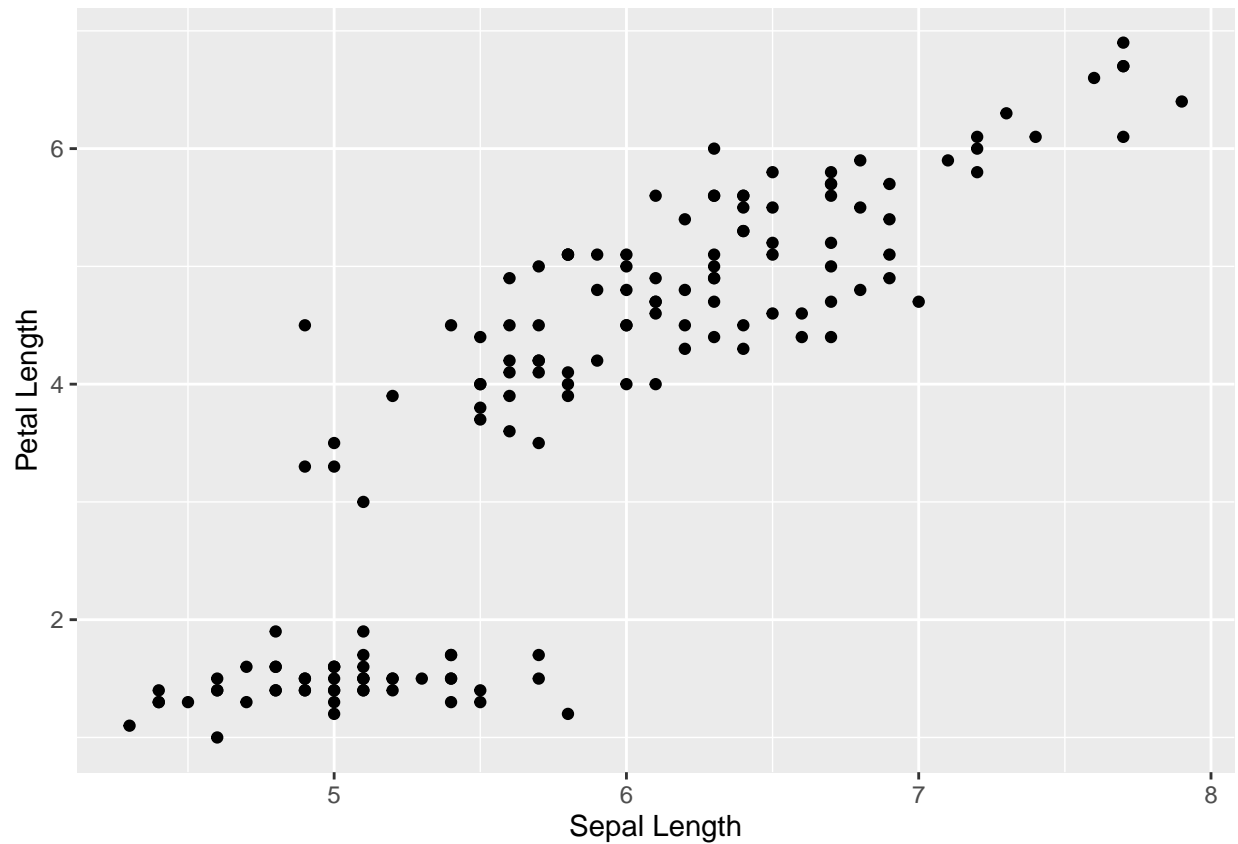
```
## 3.3
```

d) only 27% of flowers have a sepal width higher than 3.3 cm

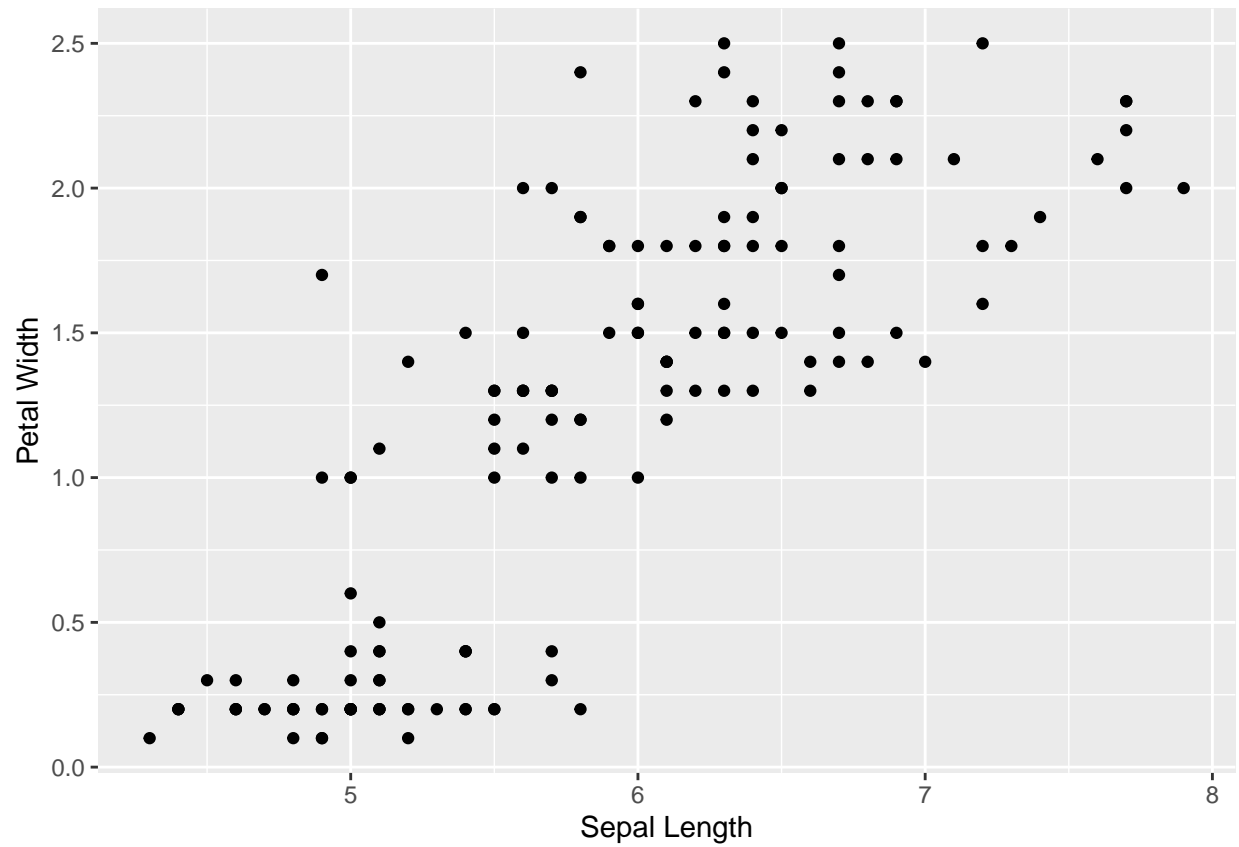
```
# Sepal Length/Sepal Width  
ggplot(iris, aes(x=Sepal.Length, y=Sepal.Width)) +  
  geom_point() +  
  labs(x="Sepal Length",  
       y="Sepal Width")
```



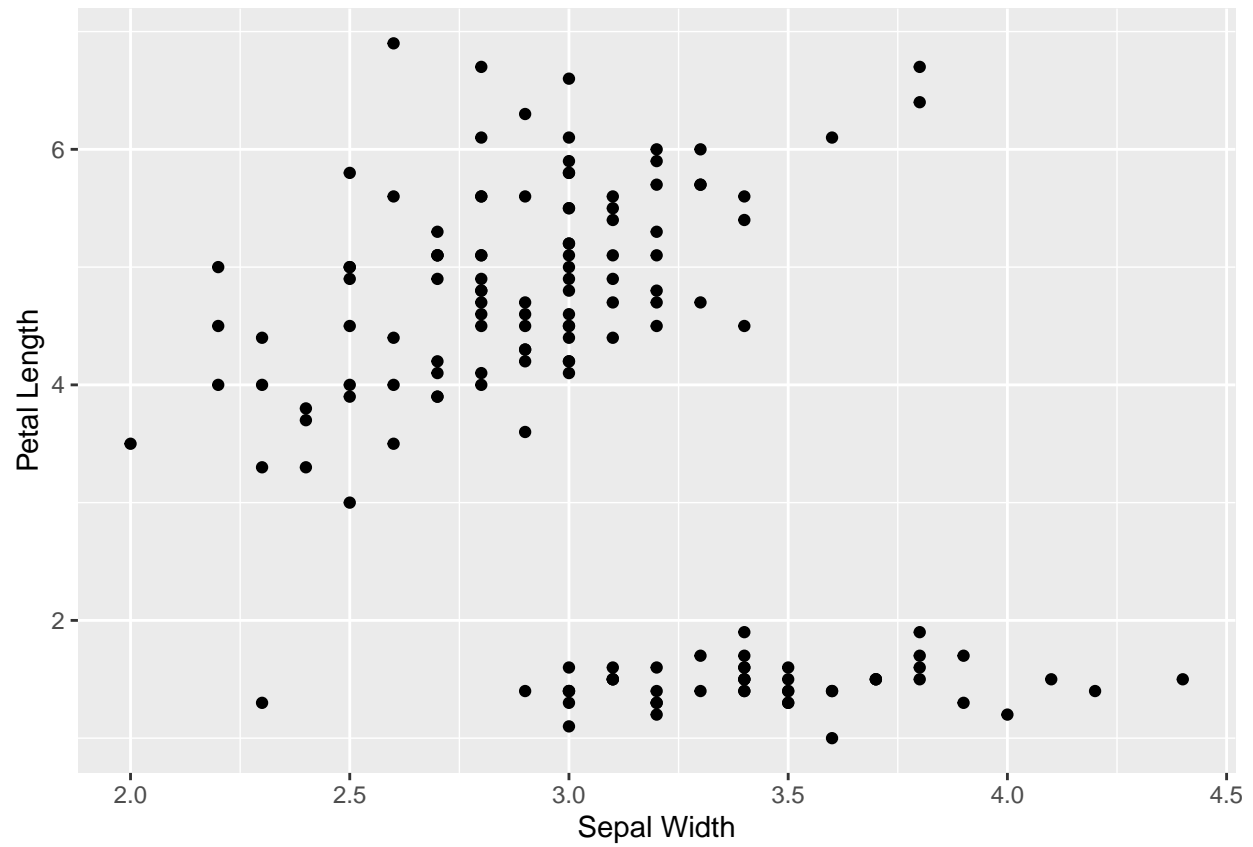
```
# Sepal Length/Petal Length
ggplot(iris, aes(x=Sepal.Length, y=Petal.Length)) +
  geom_point() +
  labs(x="Sepal Length",
       y="Petal Length")
```



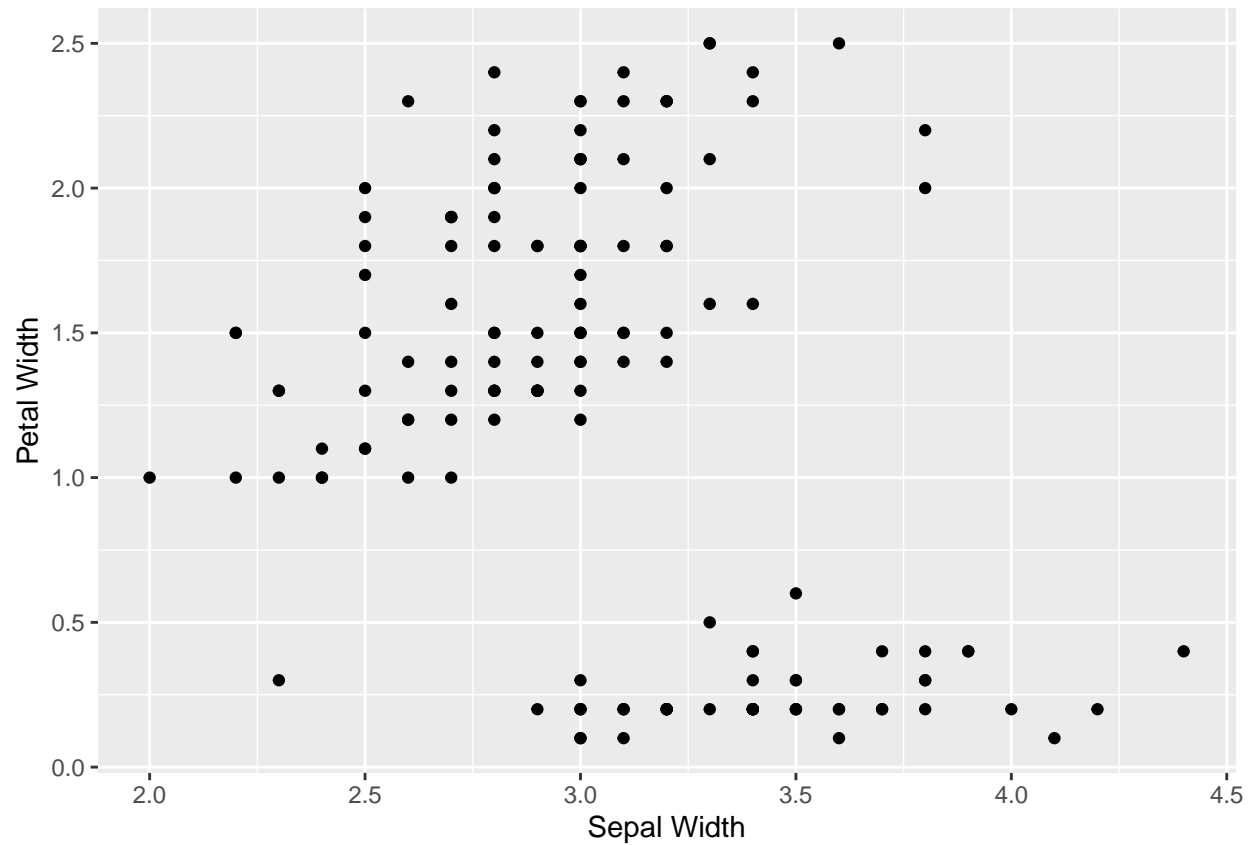
```
# Sepal Length/Petal Width
ggplot(iris, aes(x=Sepal.Length, y=Petal.Width)) +
  geom_point() +
  labs(x="Sepal Length",
       y="Petal Width")
```



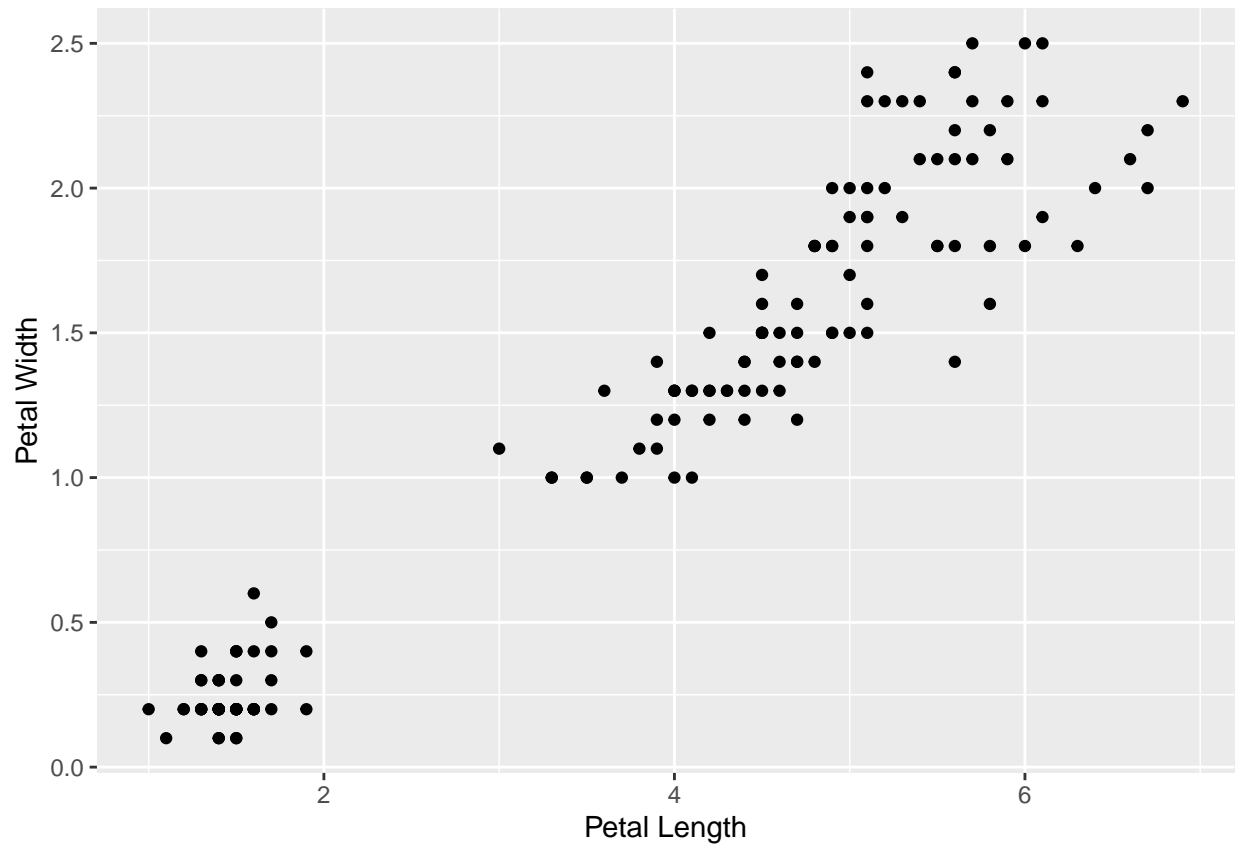
```
# Sepal Width/Petal Length  
ggplot(iris, aes(x=Sepal.Width, y=Petal.Length)) +  
  geom_point() +  
  labs(x="Sepal Width",  
       y="Petal Length")
```



```
# Sepal Width/Petal Width
ggplot(iris, aes(x=Sepal.Width, y=Petal.Width)) +
  geom_point() +
  labs(x="Sepal Width",
       y="Petal Width")
```



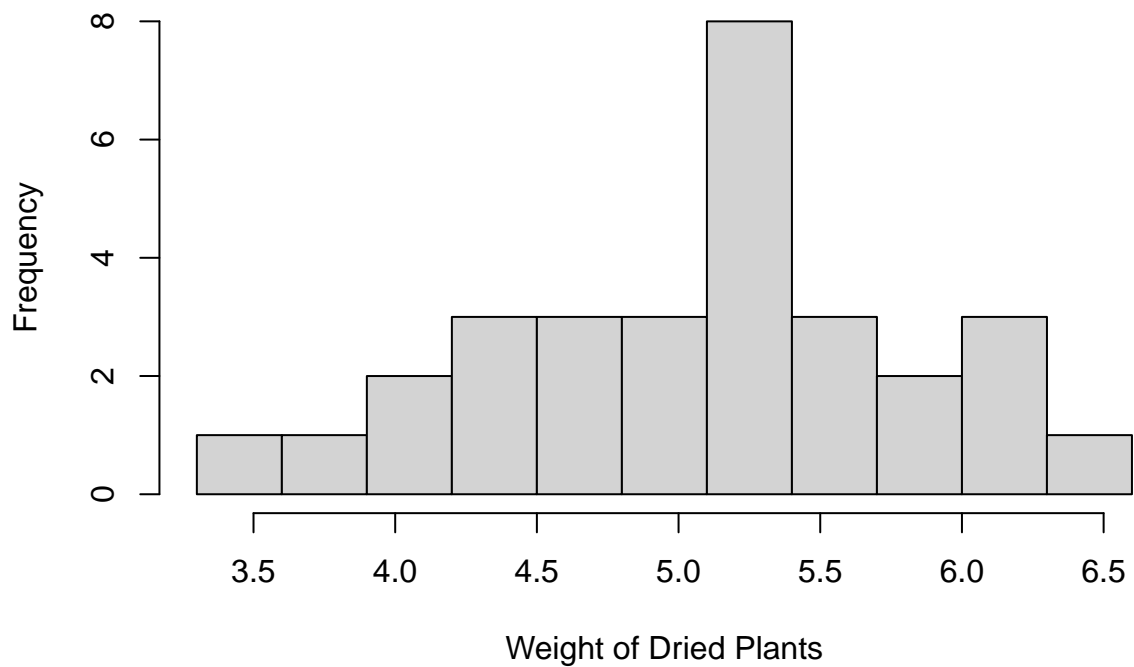
```
# Petal Length/Petal Width  
ggplot(iris, aes(x=Petal.Length, Petal.Width)) +  
  geom_point() +  
  labs(x="Petal Length",  
       y="Petal Width")
```



f) It appears that petal width and petal length have the strongest relationship, and sepal width and sepal length have the weakest relationship

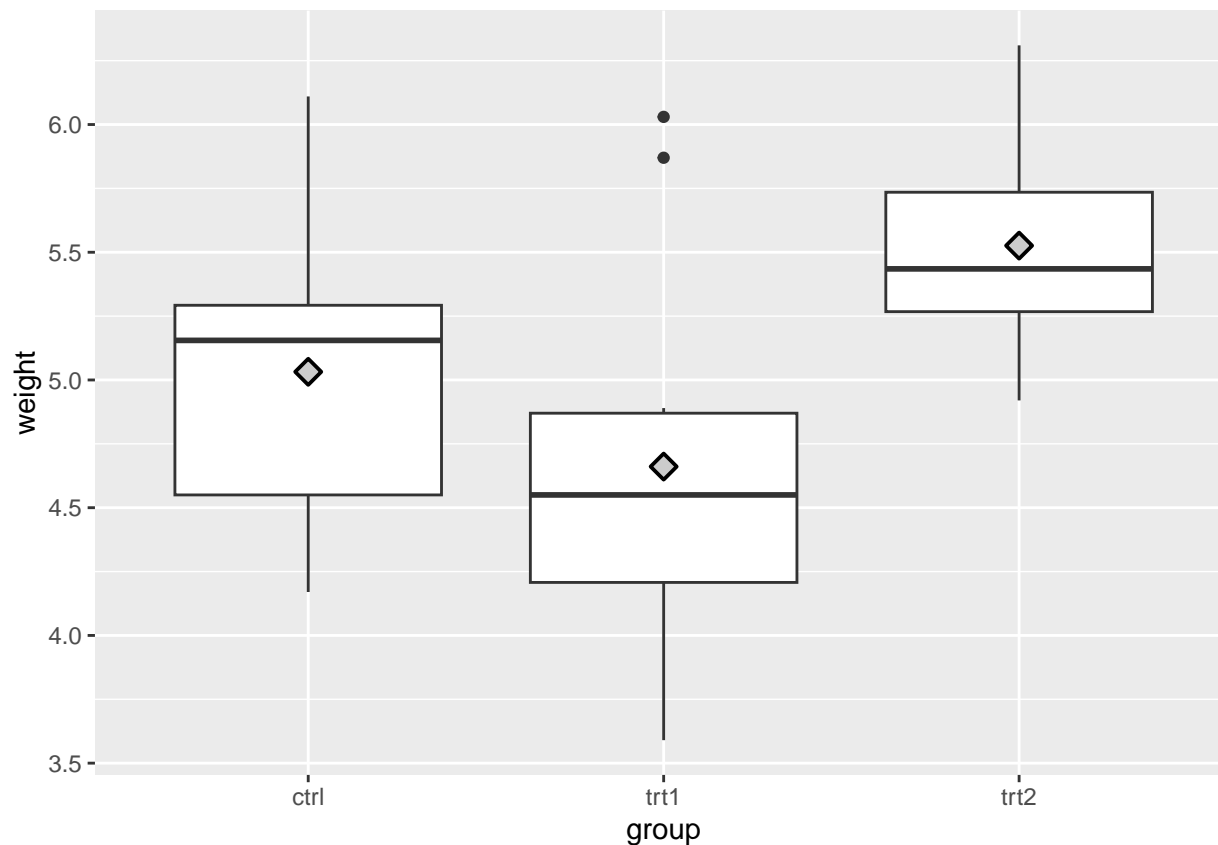
Part 2

```
hist(PlantGrowth$weight, breaks=seq(3.3,6.6,by=0.3),
     main=element_blank(),
     xlab="Weight of Dried Plants")
```



```
ggplot(PlantGrowth, aes(y=weight, x=group)) +  
  geom_boxplot() +  
  stat_summary(fun = mean, pch=23, fill = "grey80", size = 0.75)
```

```
## Warning: Removed 3 rows containing missing values or values outside the scale range  
## ('geom_segment()').
```

c) Because the trt2 boxplot has no outliers it's safe to assume the end of the 'whisker' is the minimum and because the entirety of the trt1 box is below that point it is safe to assume that at least 75% of the weights in trt1 are less than the minimum weight in trt2

```
PlantGrowth <- PlantGrowth |>
  group_by(group) |>
  mutate(Min=min(weight))
trt1 <- PlantGrowth |>
  filter(group == "trt1")

count_less <- 0
count <- 0
for (x in 1:nrow(trt1)){
  if (trt1$weight[x] < 4.92) {
    count_less <- count_less + 1}
  count <- count+1
}
(count_less/count)* 100
```

```
## [1] 80
```

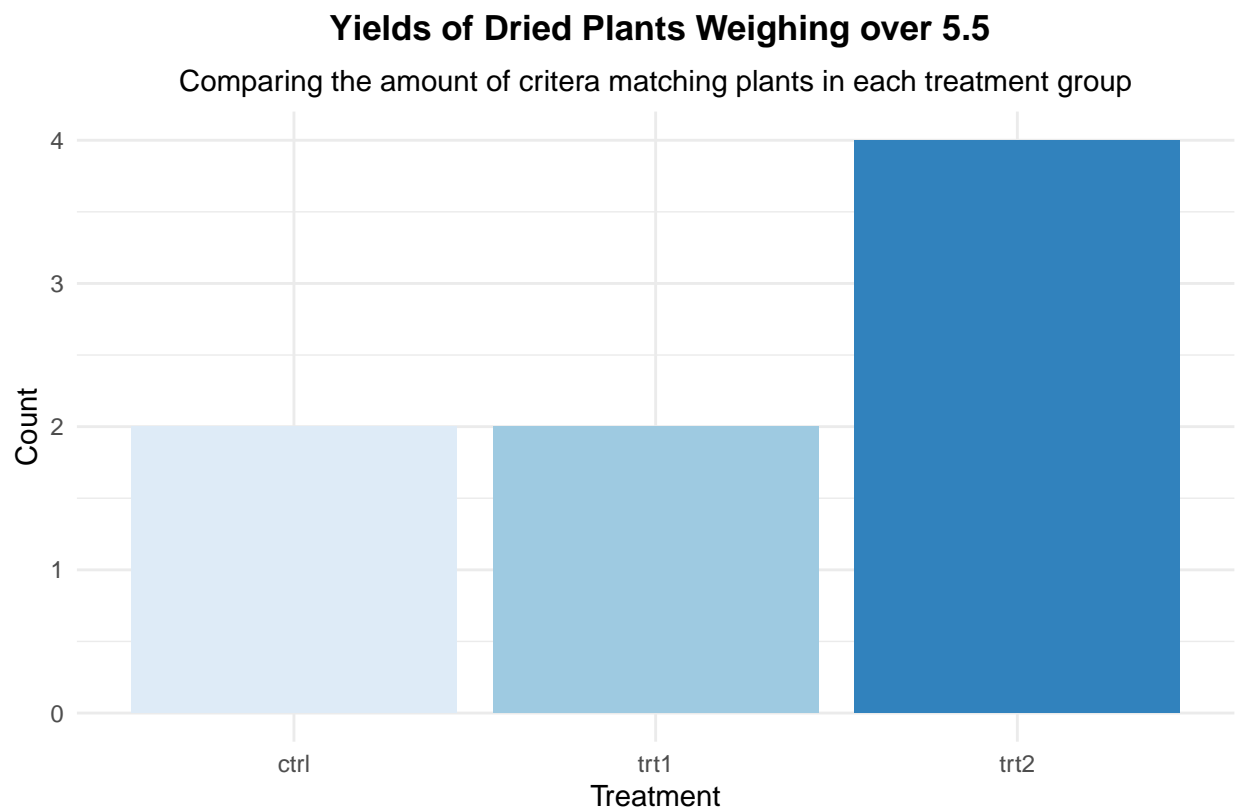
d)80% of trt1 weights are less than the minimum trt2 weight

```
PlantGrowth <- PlantGrowth |>
  ungroup() |>
  filter(weight > 5.5)
```

```

ggplot(PlantGrowth, aes(x=group, fill=group)) +
  geom_bar() +
  labs(title=~bold("Yields of Dried Plants Weighing over 5.5"),
        subtitle="Comparing the amount of criteria matching plants in each treatment group",
        caption="source: Rstudio PlantGrowth Dataset",
        x="Treatment",
        y="Count") +
  scale_fill_brewer(name="Blues") +
  theme_minimal() +
  theme(legend.position = "none",
        plot.title = element_text(hjust = 0.5),
        plot.subtitle = element_text(hjust = 0.5))

```



source: Rstudio PlantGrowth Dataset