

# lecture 3

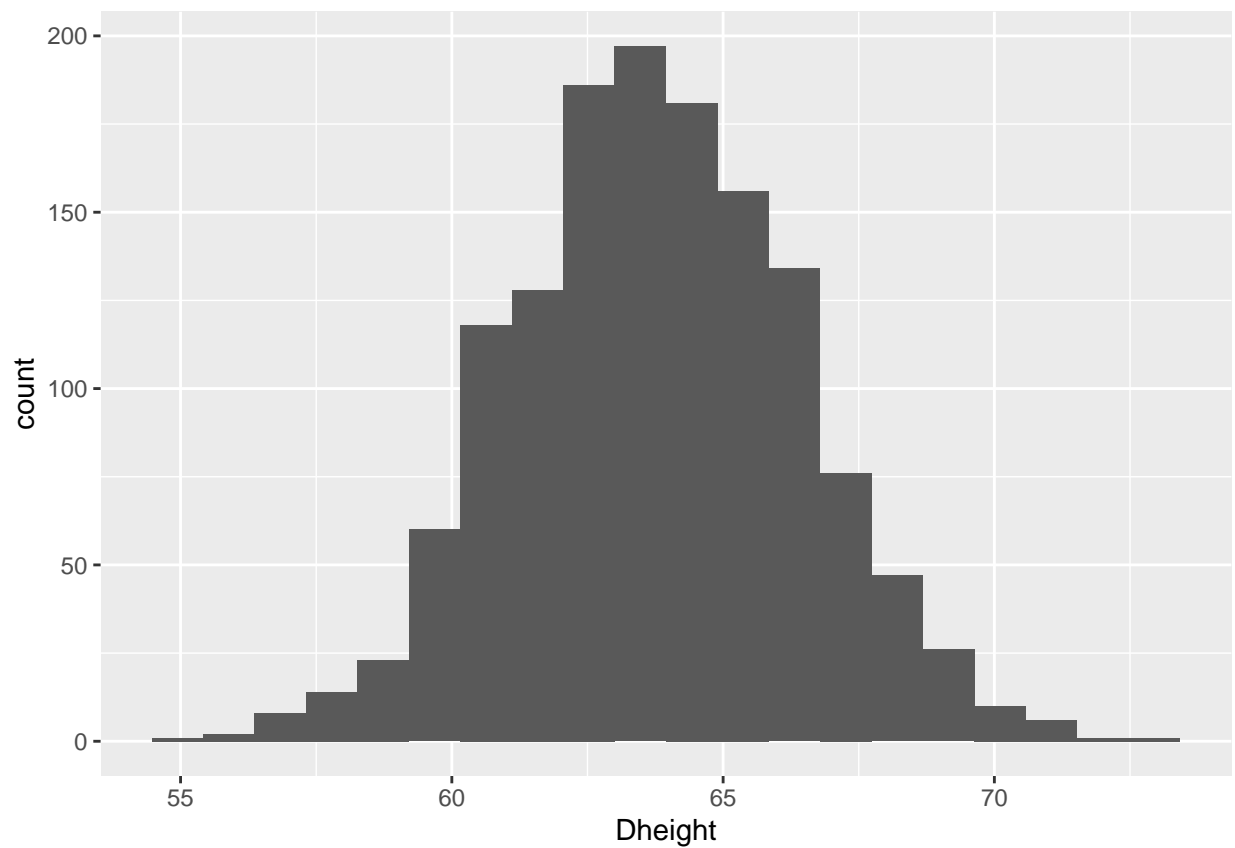
Jung-Jin Lee

1/21/2020

```
d <- read.table(file = "heights.txt", header = T, sep = " ")  
dim(d)
```

```
## [1] 1375    2
```

```
g <- ggplot(d, aes(Dheight)) + geom_histogram(bins = 20)  
print(g)
```



```
mean(d$Dheight)
```

```
## [1] 63.75105
```

```
median(d$Dheight)
```

```
## [1] 63.6
```

```
var(d$Dheight)
```

```
## [1] 6.760274
```

```
sd(d$Dheight)
```

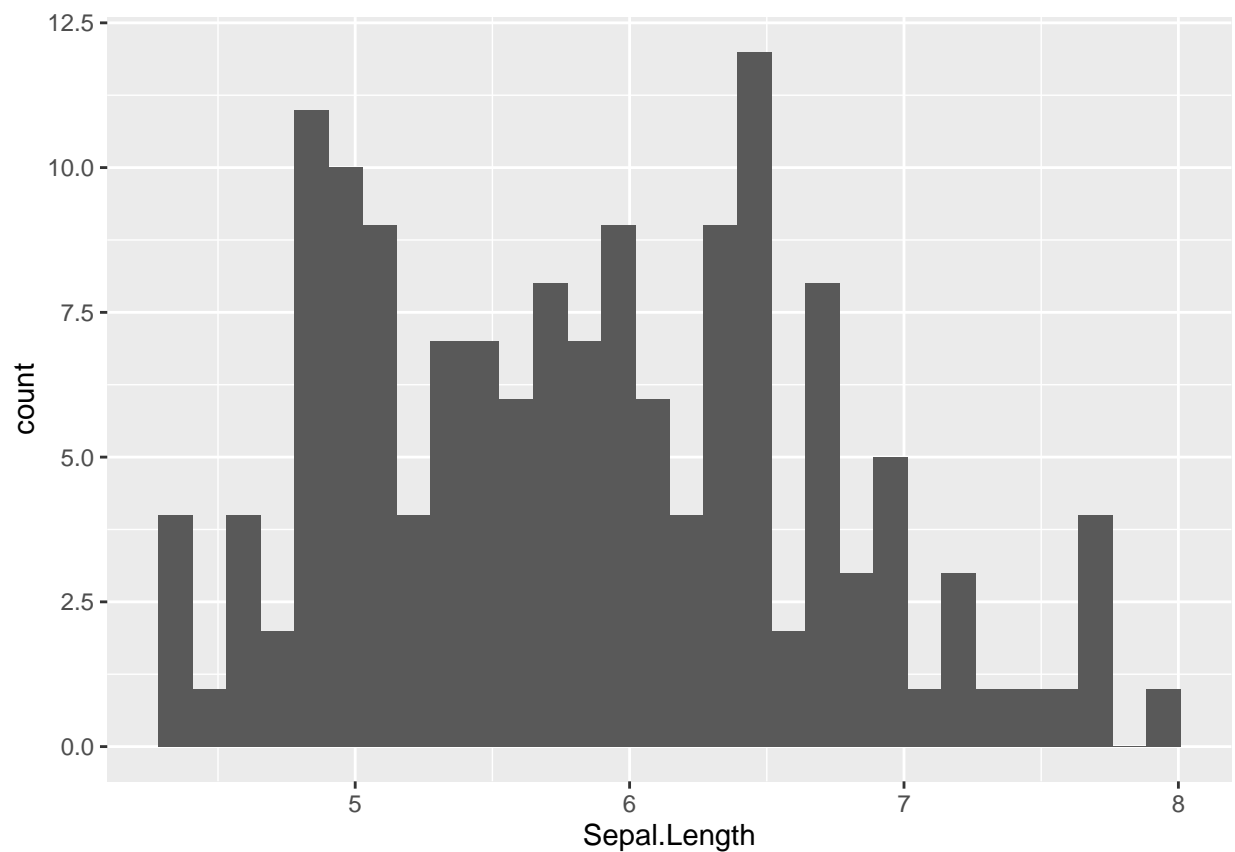
```
## [1] 2.60053
```

```
sqrt(var(d$Dheight))
```

```
## [1] 2.60053
```

```
d <- iris  
g <- ggplot(d, aes(Sepal.Length)) + geom_histogram()  
print(g)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
ggsave(filename = "histo.jpeg", g)
```

```
## Saving 6.5 x 4.5 in image  
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
f1 <- filter(d, Species == "setosa")  
dd <- filter(d, Sepal.Length >= 5.7 & Species == "virginica")
```

```
aa <- 1:5  
ifelse(aa <= 3, "what", "Ever")
```

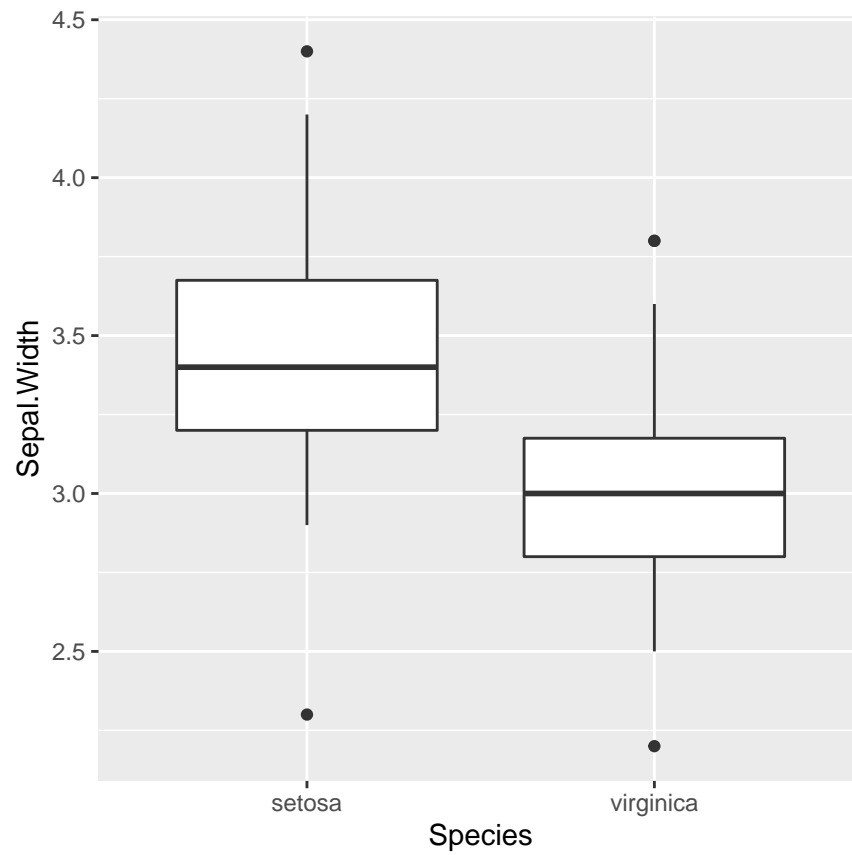
```
## [1] "what" "what" "what" "Ever" "Ever"
```

```
r1 <- mutate(d, Diff = Sepal.Length - Sepal.Width)  
r2 <- filter(r1, Diff > 1.5, Species == "setosa")  
r3 <- select(r2, Sepal.Length, Sepal.Width, Diff)  
r4 <- arrange(r3, desc(Diff))
```

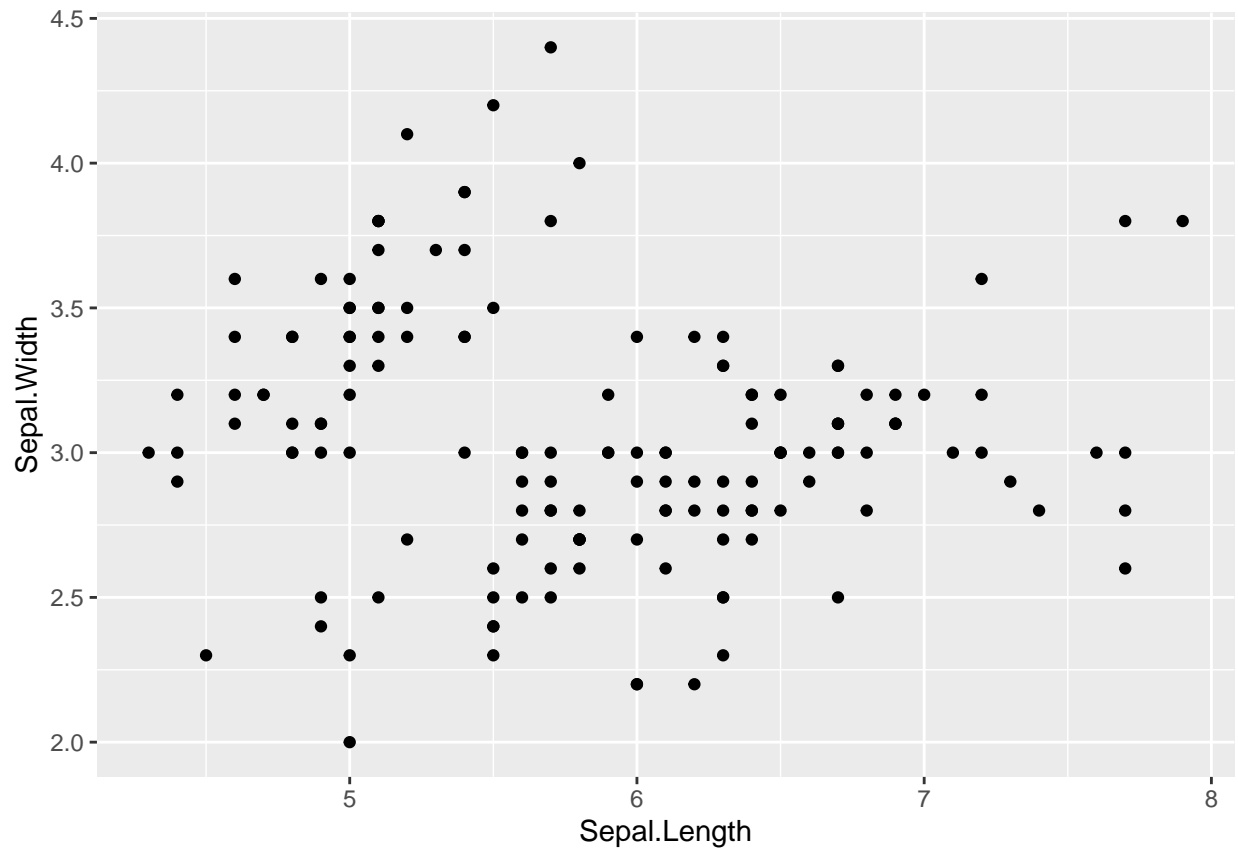
```
cal_wide <- data.frame(subject = paste0("Patient", 1:4),  
                       breakfast = c(300, 350, 300, 250),  
                       lunch = c(500, 500, 500, 450),  
                       dinner = c(600, 650, 450, 500))  
print(cal_wide)
```

```
##      subject breakfast lunch dinner  
## 1 Patient1      300    500    600  
## 2 Patient2      350    500    650  
## 3 Patient3      300    500    450  
## 4 Patient4      250    450    500
```

```
d %>%  
  filter(Species %in% c("setosa", "virginica")) %>%  
  ggplot(aes(Species, Sepal.Width)) +  
  geom_boxplot() +  
  theme(aspect.ratio = 1)
```

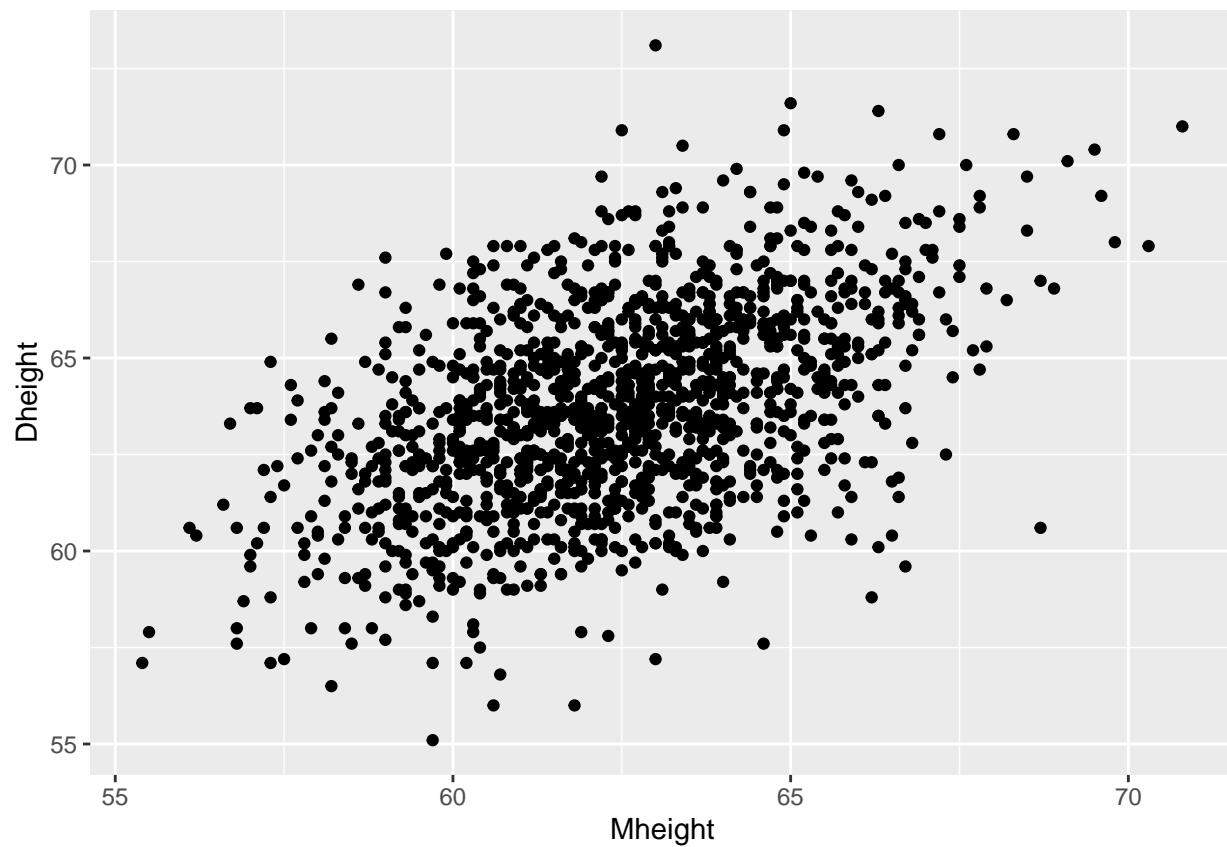


```
g <- ggplot(d, aes(x = Sepal.Length, y = Sepal.Width)) +  
  geom_point()  
print(g)
```



```
g2 <- g + theme(aspect.ratio = 1)
```

```
"heights.txt" %>%  
  read.table(header = T, sep = " ") %>%  
  ggplot(aes(Mheight, Dheight)) +  
  geom_point()
```



3 | 0

```
## [1] TRUE
```