Appendix

Jadon Fowler, STA 570 Section 1

Homework 9, 2023/04/20

```r
library(car)
library(ggplot2)
library(dplyr)
library(ggfortify)
library(emmeans)
library(multcompView)
library(multcomp)
```

```r
# 1
pf(29.466, df1=3, df2=13)
```

```
## [1] 0.9999954
```

```r
# 2
pf(.3458652541,df1=2,df2=6)
```

```
## [1] 0.2791608
```

```r
# 3
pf(31.147,df1=2,df2=9)
```

```
## [1] 0.9999098
```

```r
data.3 <- data.frame(grp = rep(c('1','2','3'), each = 4 ), y = c(4,6,6,8,8,8,6,6,12,13,15,16))
model <- lm(y ~ grp, data=data.3)
anova(model)
```

```
## Analysis of Variance Table
##
## Response: y
##            Df Sum Sq Mean Sq F value    Pr(>F)
## grp         2    152  76.000  31.091 9.087e-05 ***
## Residuals   9     22   2.444
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# 5
Levi <- read.csv('https://raw.githubusercontent.com/dereksonderegger/570/master/data-raw/Levi.csv')
ggplot(Levi) +
  geom_boxplot(aes(x=Plant, y=Waste))
```
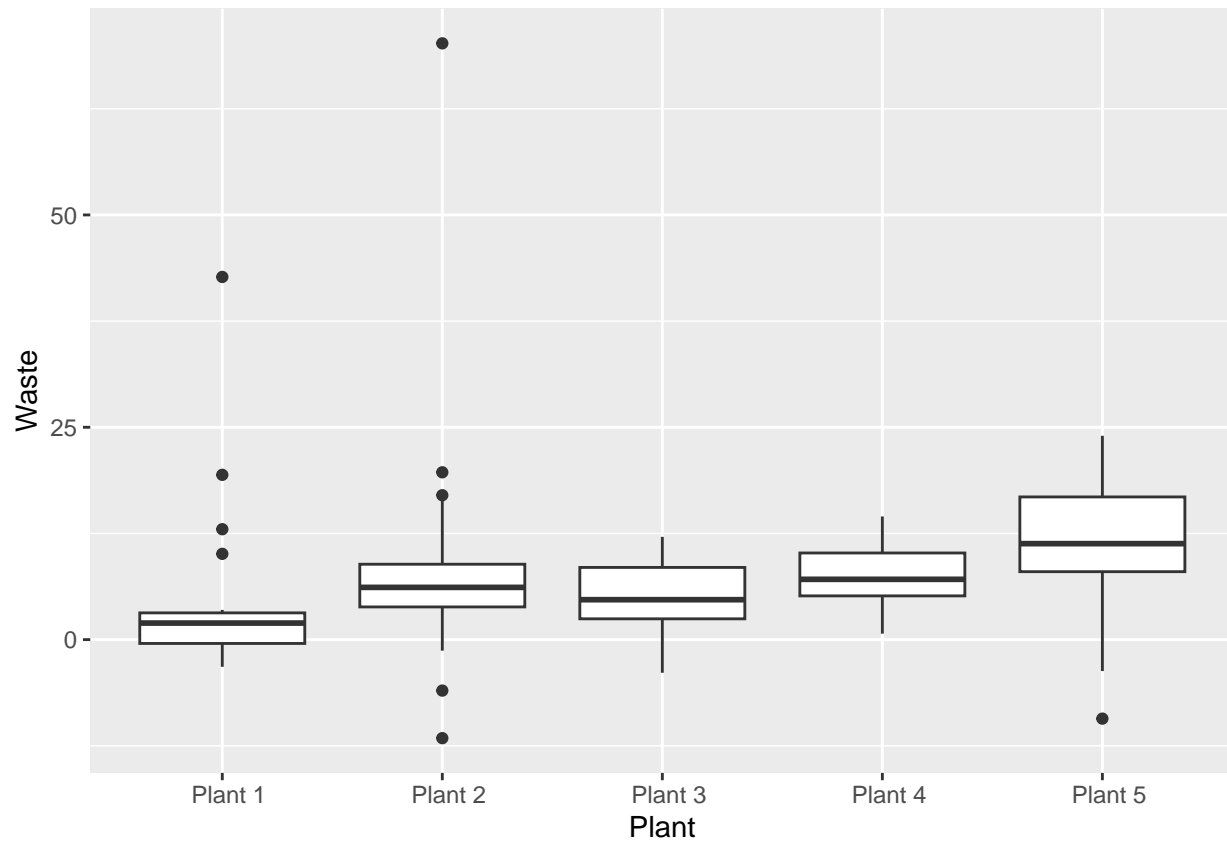
```
car::leveneTest(Waste~Plant, Levi)
```

```
## Warning in leveneTest.default(y = y, group = group, ...): group coerced to
## factor.
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  4  1.2235 0.3065
##       90
```

```
levi.model <- lm(Waste ~ Plant, data=Levi)
anova(levi.model)
```
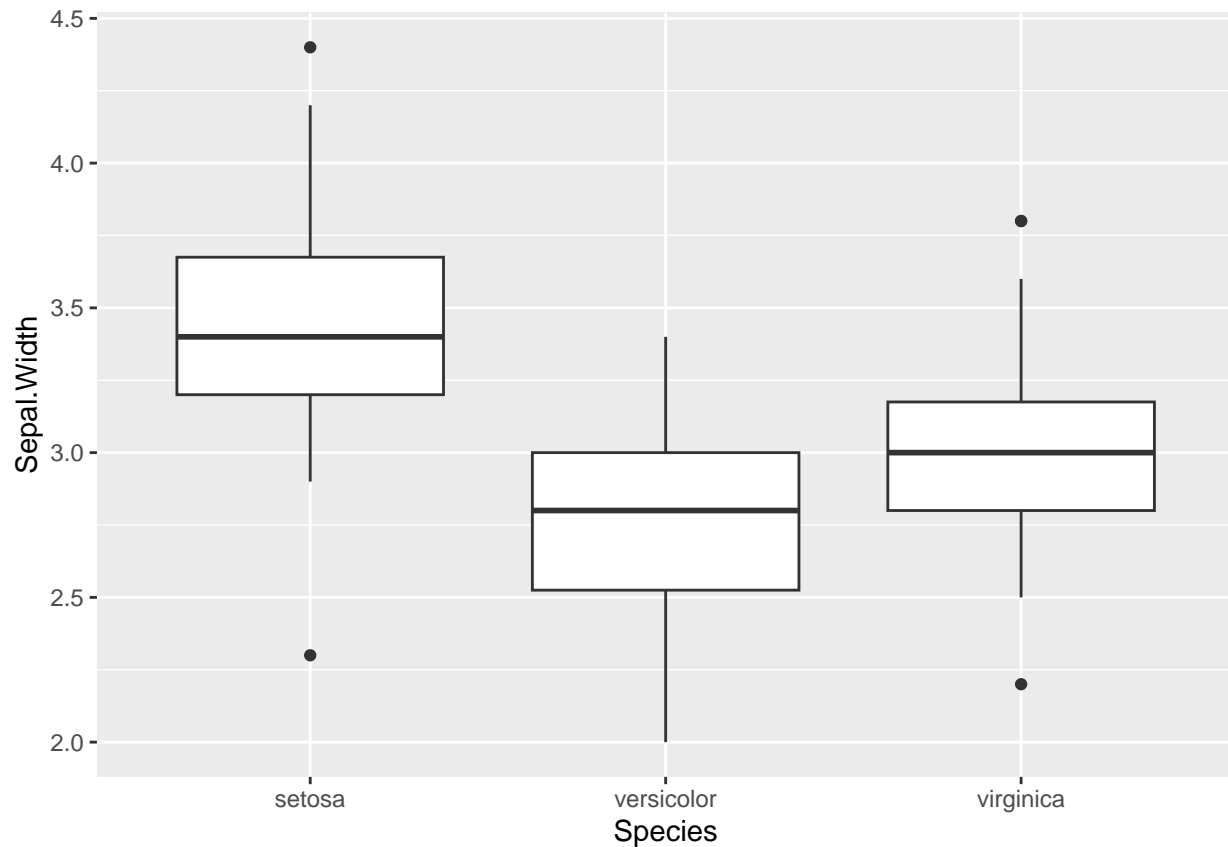
```
## Analysis of Variance Table
##
## Response: Waste
##           Df Sum Sq Mean Sq F value Pr(>F)
## Plant      4  450.9 112.730  1.1596  0.334
## Residuals 90 8749.1  97.212
```

```
shapiro.test(resid(levi.model))
```

```
##
```

```
##   Shapiro-Wilk normality test
##
## data:  resid(levi.model)
## W = 0.71955, p-value = 3.529e-12
```

```r
# 6
data('iris')
ggplot(iris) +
  geom_boxplot(aes(x=Species,y=Sepal.Width))
```



```r
#b
car::leveneTest(Sepal.Width~Species, iris)
```
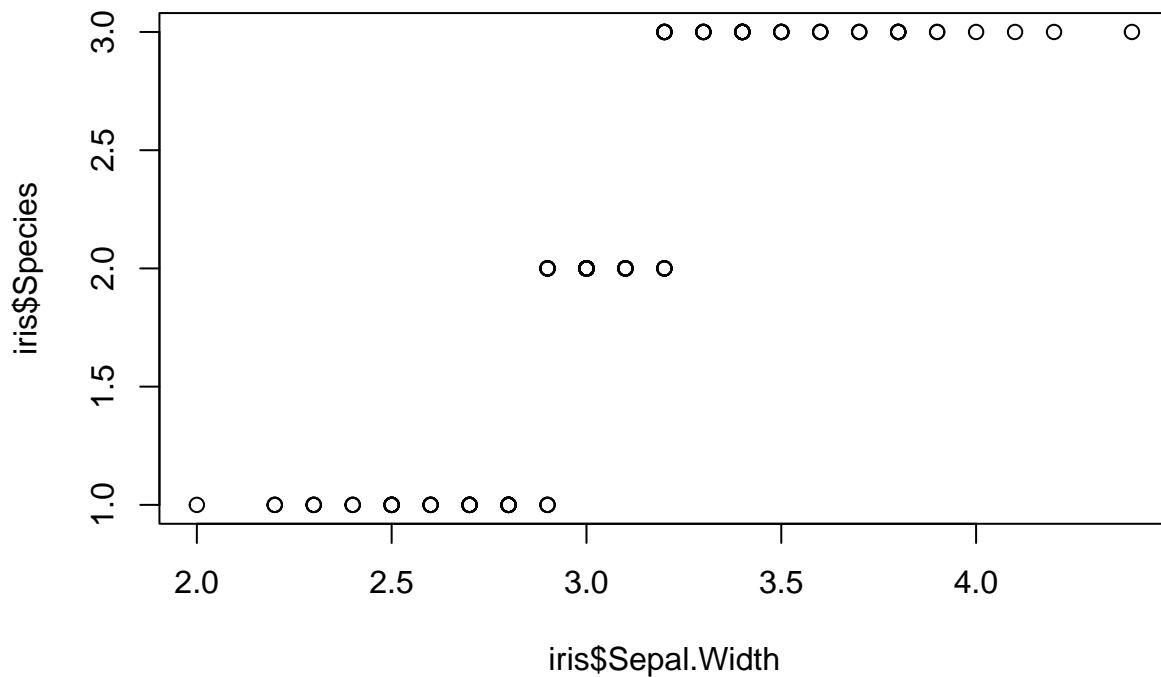
```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   2  0.5902 0.5555
##       147
```

```r
#c
iris.model <- lm(Sepal.Width ~ Species, data=iris)
anova(iris.model)
```

```
## Analysis of Variance Table
##
```

```
## Response: Sepal.Width
##            Df Sum Sq Mean Sq F value    Pr(>F)
## Species     2 11.345  5.6725   49.16 < 2.2e-16 ***
## Residuals 147 16.962  0.1154
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqplot(iris$Sepal.Width, iris$Species)
```



```
shapiro.test(resid(iris.model))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid(iris.model)
## W = 0.98948, p-value = 0.323
```

```
#e
emmeans::emmeans(iris.model, pairwise ~ Species)
```

```
## $emmeans
##  Species     emmean    SE  df lower.CL upper.CL
##  setosa        3.43 0.048 147     3.33     3.52
##  versicolor    2.77 0.048 147     2.68     2.86
```

```
##  virginica     2.97 0.048 147     2.88     3.07
##
## Confidence level used: 0.95
##
## $contrasts
##  contrast                estimate     SE  df t.ratio p.value
##  setosa - versicolor        0.658 0.0679 147   9.685  <.0001
##  setosa - virginica         0.454 0.0679 147   6.683  <.0001
##  versicolor - virginica    -0.204 0.0679 147  -3.003  0.0088
##
## P value adjustment: tukey method for comparing a family of 3 estimates
```