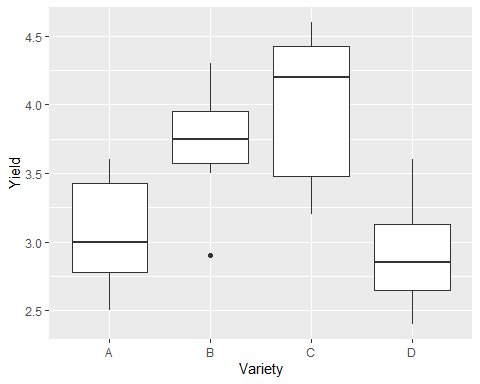
STAR511: HW 6

Megan Sears

# Q1

\*\*



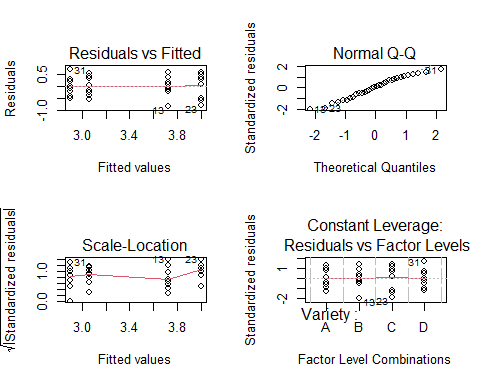
# Q2

A. The null hypothesis is that the mean yields for the varieties are equal, and the alternative hypothesis is that one or more difference exists between the mean yield for the varieties.

## Analysis of Variance Table  
##   
## Response: Yield  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Variety 3 6.6209 2.20698 11.047 5.85e-05 \*\*\*  
## Residuals 28 5.5938 0.19978   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

C. The p-value is 0.0000585. Here we will reject the null hypothesis since the p-value is less than alpha\*\*. This indicates that there is strong evidence indicating differences among the popuation-level mean yields for varieties.

# Q3



# Q4

## contrast estimate SE df t.ratio p.value  
## A - B -0.662 0.223 28 -2.964 0.0061  
## A - C -0.938 0.223 28 -4.195 0.0002  
## A - D 0.163 0.223 28 0.727 0.4732  
## B - C -0.275 0.223 28 -1.231 0.2287  
## B - D 0.825 0.223 28 3.692 0.0010  
## C - D 1.100 0.223 28 4.922 <.0001

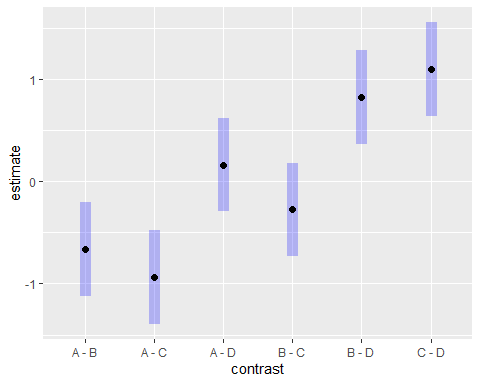
Groups B&C and A&D have p-values greater than 0.05, which indicates the null for those groups should not be rejected and the means are not significantly different. The other groups (A&B, A&C, B&D, C&D) have p-values less than 0.05, which indicates the null for those should be rejected and the means are significantly different.

#Q5

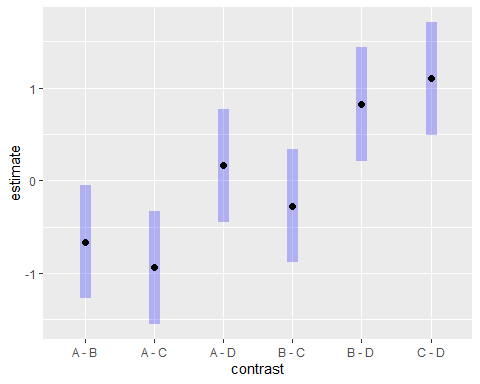
## contrast estimate SE df t.ratio p.value  
## A - B -0.662 0.223 28 -2.964 0.0295  
## A - C -0.938 0.223 28 -4.195 0.0013  
## A - D 0.163 0.223 28 0.727 0.8854  
## B - C -0.275 0.223 28 -1.231 0.6132  
## B - D 0.825 0.223 28 3.692 0.0050  
## C - D 1.100 0.223 28 4.922 0.0002  
##   
## P value adjustment: tukey method for comparing a family of 4 estimates

This produced the same conclusions as above.Groups B&C and A&D have p-values greater than 0.05, which indicates the null for those groups should not be rejected and the means are not significantly different. The other groups (A&B, A&C, B&D, C&D) have p-values less than 0.05, which indicates the null for those should be rejected and the means are significantly different.

# Q6



# Q7

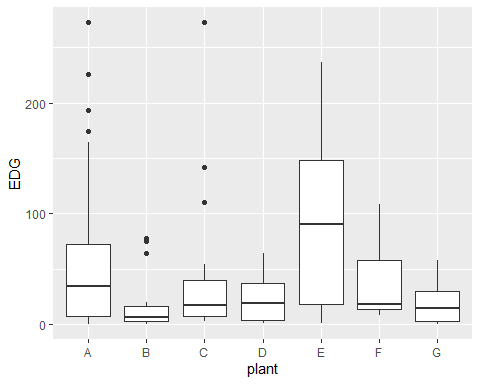


When comparing the 2 plots, I am able to tell that the plot in Q7 has had a multiple comparisons adjustment because it has wider confidence intervals (length of blue bars) than the plot in Q6.

# Q8

the means and the confidence intervals \*\*\*\*\*\*\*

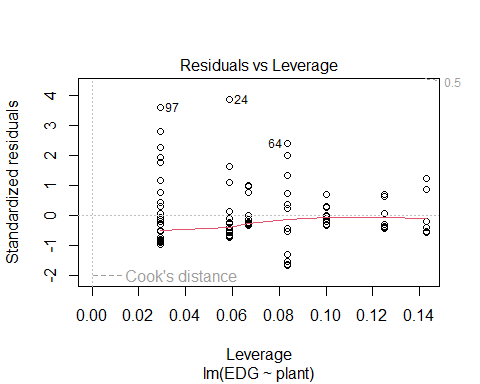
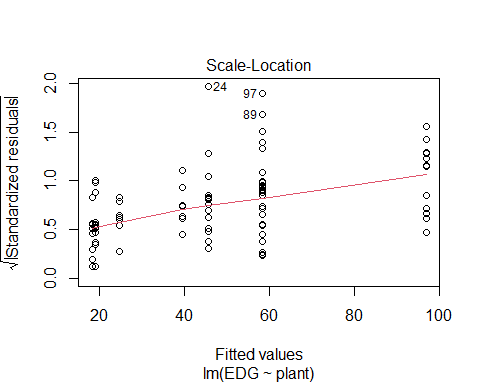
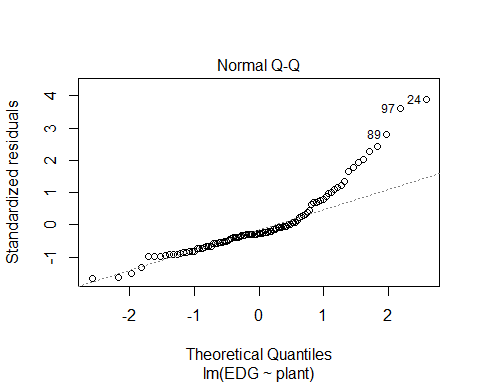
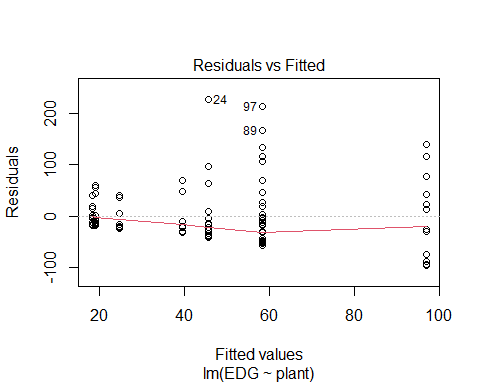
#Q9



# Q10

## Analysis of Variance Table  
##   
## Response: EDG  
## Df Sum Sq Mean Sq F value Pr(>F)   
## plant 6 58745 9790.9 2.6761 0.01912 \*  
## Residuals 96 351233 3658.7   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Q11



# Appendix

#Retain (and do not edit) this code chunk!!!  
library(knitr)  
knitr::opts\_chunk$set(echo = FALSE)  
knitr::opts\_chunk$set(message = FALSE)  
  
library(tidyverse)  
library(here)  
  
#Q1  
corn <- read\_csv(here('./HW6/CornYield.csv'))  
  
ggplot(corn, aes(x=Variety, y=Yield)) +  
 geom\_boxplot()  
   
#Q2  
corn\_anov <- lm(Yield ~ Variety, data = corn)  
  
anova(corn\_anov)  
  
#Q3  
par(mfrow=c(2,2))  
plot(lm(Yield ~ Variety, data = corn))  
  
library('emmeans')  
  
#Q4  
lmfit <- lm(Yield ~ Variety, data = corn)  
emout <- emmeans(lmfit, ~ Variety)  
pairs(emout, adjust = 'none')  
  
#Q5  
pairs(emout)  
  
#Q6  
  
plot(pairs(emout, adjust = 'none'), horizontal = F)  
  
#Q7  
plot(pairs(emout), horizontal = F)  
  
#Q9  
pp <- read\_csv(here('./HW6/ex8-23.txt'), quote = "'")  
  
pp <- pp %>%  
 pivot\_longer(everything(), names\_to = 'plant', values\_to = 'EDG') %>%  
 mutate(plant = as\_factor(plant)) %>%  
 na.omit()  
  
ggplot(pp, aes(x = plant, y = EDG)) +  
 geom\_boxplot()  
  
  
fit <- lm(EDG ~ plant, data = pp)  
anova(fit)  
  
  
#Q11  
plot(fit)