HW5

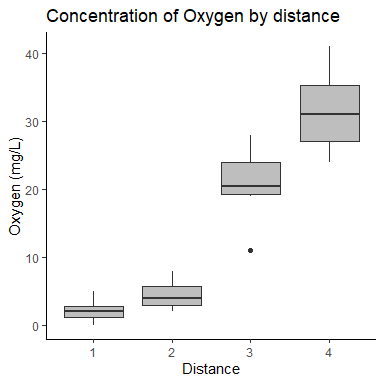
Henrique Magalhaes Rio

# Question 1

R is treating the groups as numbers instead of treating as a as factor. it could be fixed by first transforming the data using as.factor

# Question 2

## Part 1 (A)



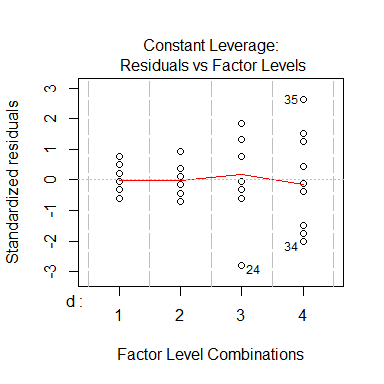
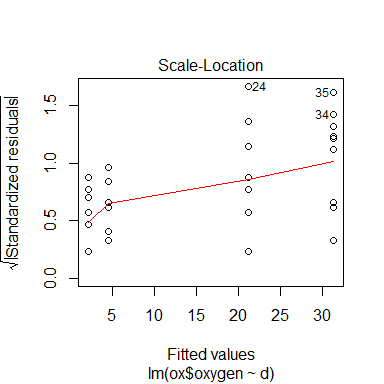
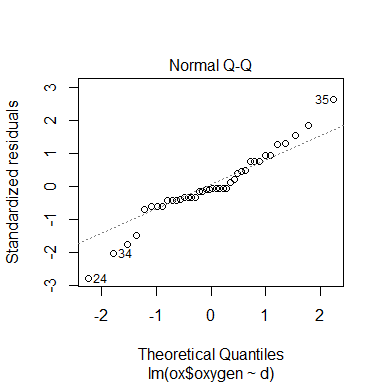
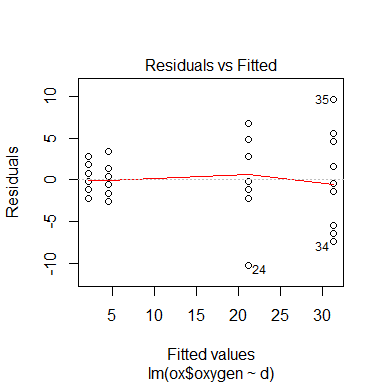
## Part 2 (B)

## Analysis of Variance Table  
##   
## Response: ox$oxygen  
## Df Sum Sq Mean Sq F value Pr(>F)   
## d 3 5793.1 1931.03 129.7 < 2.2e-16 \*\*\*  
## Residuals 36 536.0 14.89   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Part 3 (C)

We reject the null hyphoteses that the average concentration of oxygen between all distances are equal, and can conclude that at least one of them differs. (P-value<0.0001)

## Part 4 (C)



## Part 5 (D)

Looking at the QQ plot, we can conclude that the assumption of normality of has been violated due the the “S” shaped graph. However, despite a light cone shape on the residuals vs fitted graph i do not believe to be enough to discard the equality of variance assumption.

# Question 3

95%CI for control= (91.78825,97.21175)

# Question 4

## [1] 3.834997

# Question 5

we fail to reject the null hyphoteses that the there is no difference between the average growth of the control group and the ABA treatment. 95%CIcontrol=(91.78825,97.21175)

# Question 6

## [1] 5.045475

# Question 7

## [1] 89.45453 99.54547

95%CIcontrol=(89.45453,99.54547)

# Question 8

we fail to reject the null hyphoteses that the there is no difference between the average growth of the control group and the ABA treatment. 95%CIcontrol=(89.45453,99.54547)

# Question 9

the coverage is 0.7

# Appendix

library(knitr)  
library(ggplot2)  
library(dplyr)  
library(tidyverse)  
library(broom)  
library(splines)  
library(caret)  
library(Matrix)  
library(MASS)  
library(formatR)  
knitr::opts\_chunk$set(echo = FALSE, message = FALSE, warning = FALSE, fig.width = 4,   
 fig.height = 4, tidy = TRUE)  
  
# Q2A  
ox <- read\_csv("oxygen.csv")  
  
  
d <- as.factor(ox$distance)  
  
  
  
ggplot(data = ox, aes(x = d, y = ox$oxygen)) + geom\_boxplot(fill = "gray") +   
 labs(title = "Concentration of Oxygen by distance ", x = "Distance", y = "Oxygen (mg/L)") +   
 theme\_classic()  
  
  
  
# q2B  
lmout <- lm(ox$oxygen ~ d)  
anova(lmout)  
plot(lmout)  
  
# Q3  
tcrit <- qt(0.975, df = 20)  
  
ci <- 94.5 + c(-1, 1) \* tcrit \* sqrt(10.14/6)  
# Q4  
LSD <- tcrit \* sqrt(2 \* 10.14/6)  
  
LSD  
  
# q6  
lsdb <- qt(1 - 0.05/(2 \* 4), df = 20) \* sqrt(2 \* 10.14/6)  
lsdb  
lsdb <- qt(1 - 0.05/(2 \* 4), df = 20) \* sqrt(2 \* 10.14/6)  
ci <- 94.5 + c(-1, 1) \* lsdb  
ci