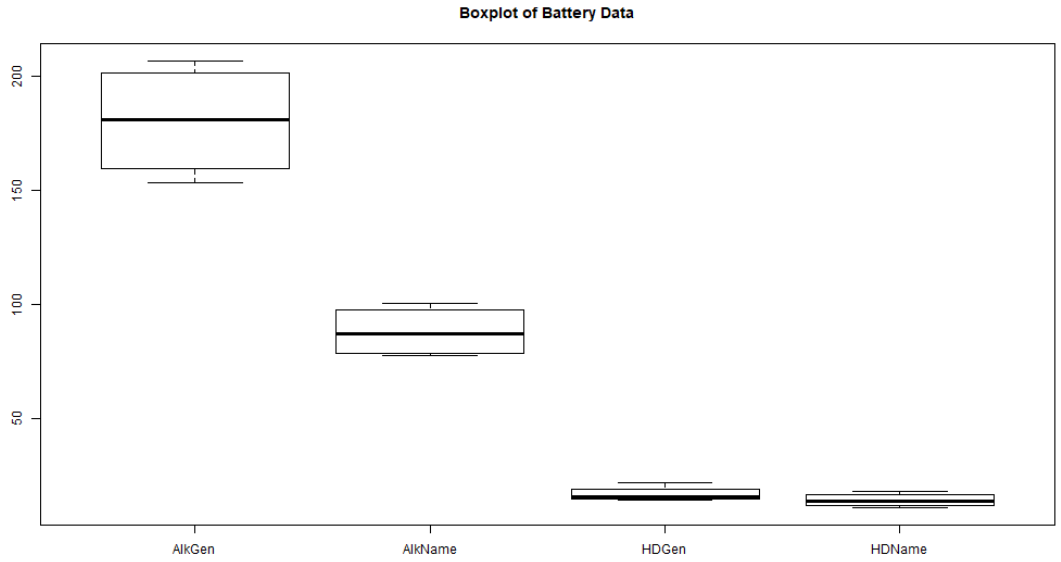
STAT 461: Homework 6

Name: Kyle Salitrik | ***ID#:*** *997543474* | ***PSU ID:*** *kps168*

October 17, 2018

PROBLEM 1


Boxplot of Battery Data
200
150
100
50
AlkGen
AlkName
HDGen
HDName

PROBLEM 2

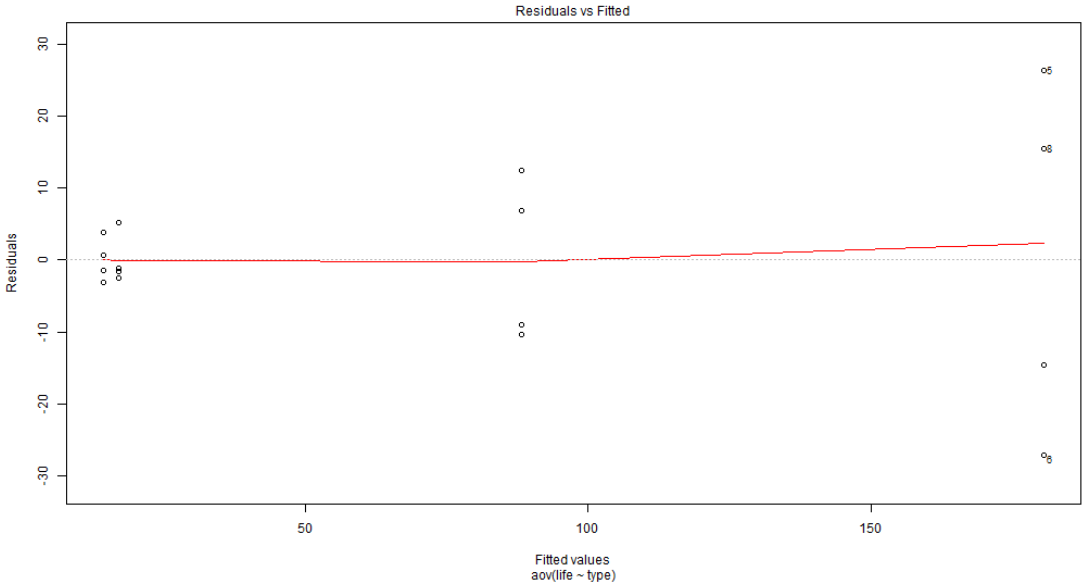
a)

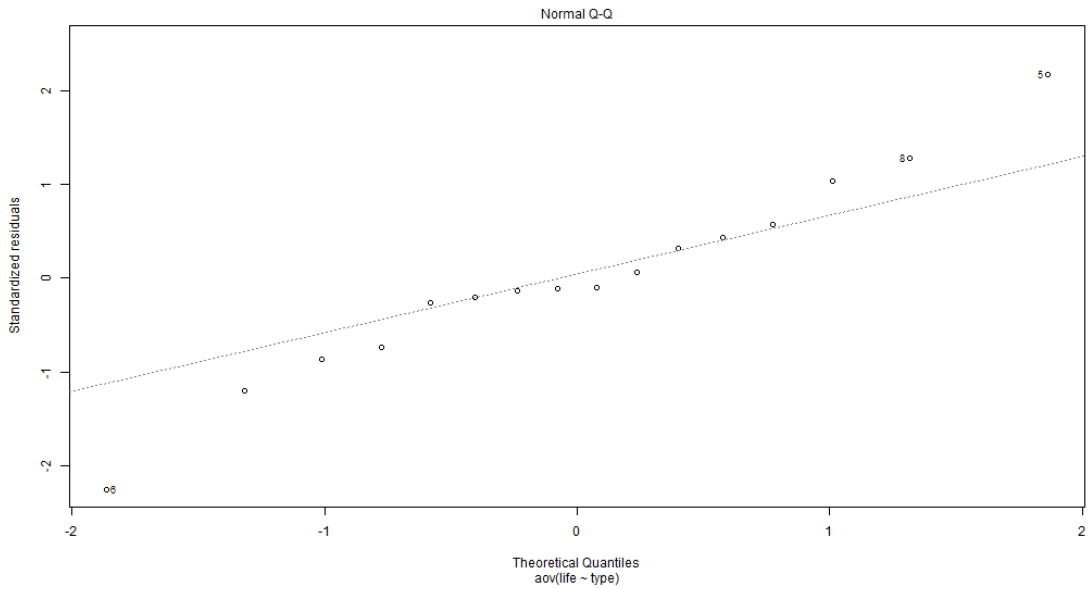
$$Y_{it} = \mu + \tau_i + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma^2)$$

$i = \text{AlkName}, \text{AlkGen}, \text{HDName}, \text{HDGen}$

$t = 1, \ldots, r_i; \quad r_{AlkName} = r_{AlkGen} = r_{HDName} = r_{HDGen} = 4$

b)





c)

No. The Residuals vs Fitted plots clearly show an increasing trend in the residuals.

PROBLEM 2

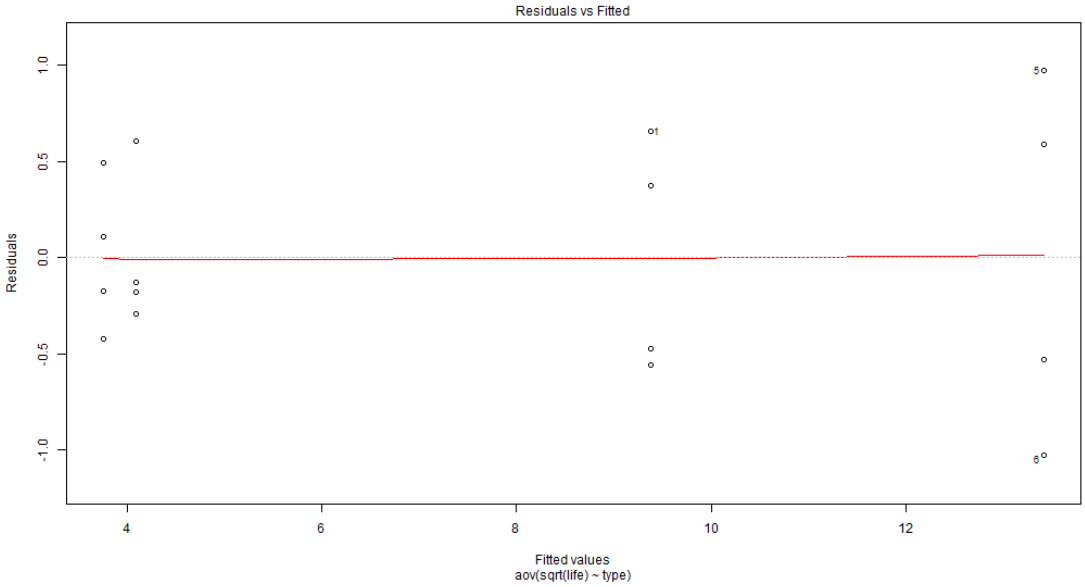
a)

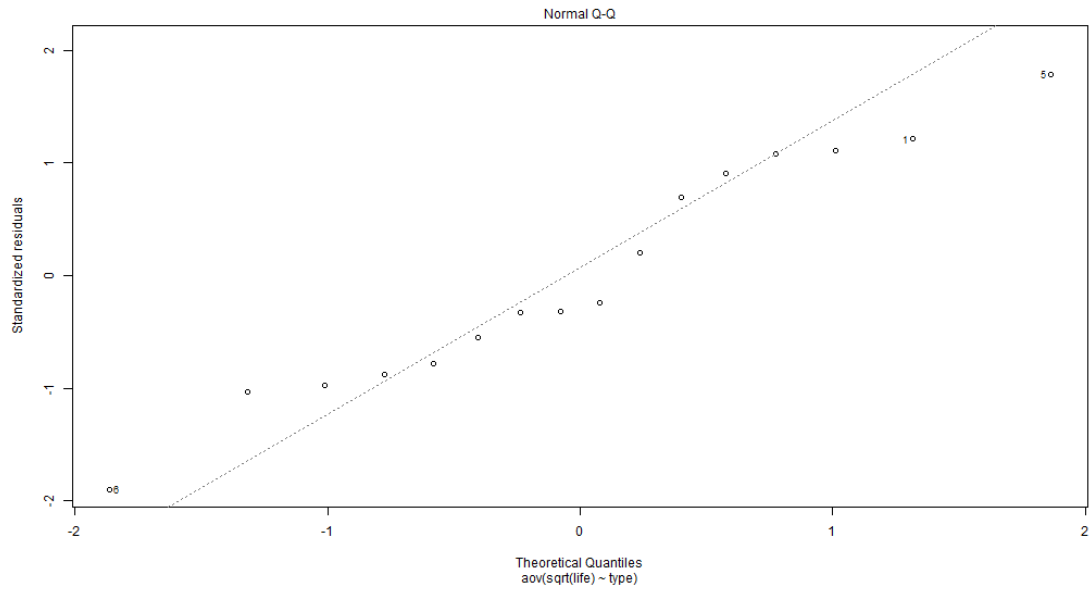
$$\sqrt{Y_{it}} = \mu + \tau_i + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma^2)$$

$i = \text{AlkName}, \text{AlkGen}, \text{HDName}, \text{HDGen}$

$t = 1, \ldots, r_i; \quad r_{AlkName} = r_{AlkGen} = r_{HDName} = r_{HDGen} = 4$

b)





c)

The Residuals vs Fitted plot is better than the original model, but there is still somewhat of a trend.

PROBLEM 4

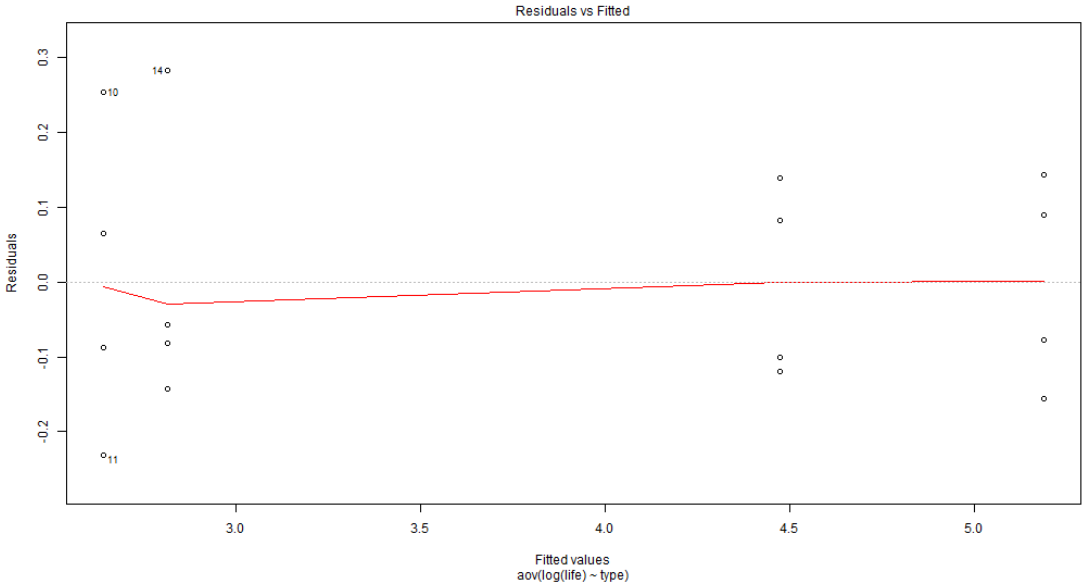
a)

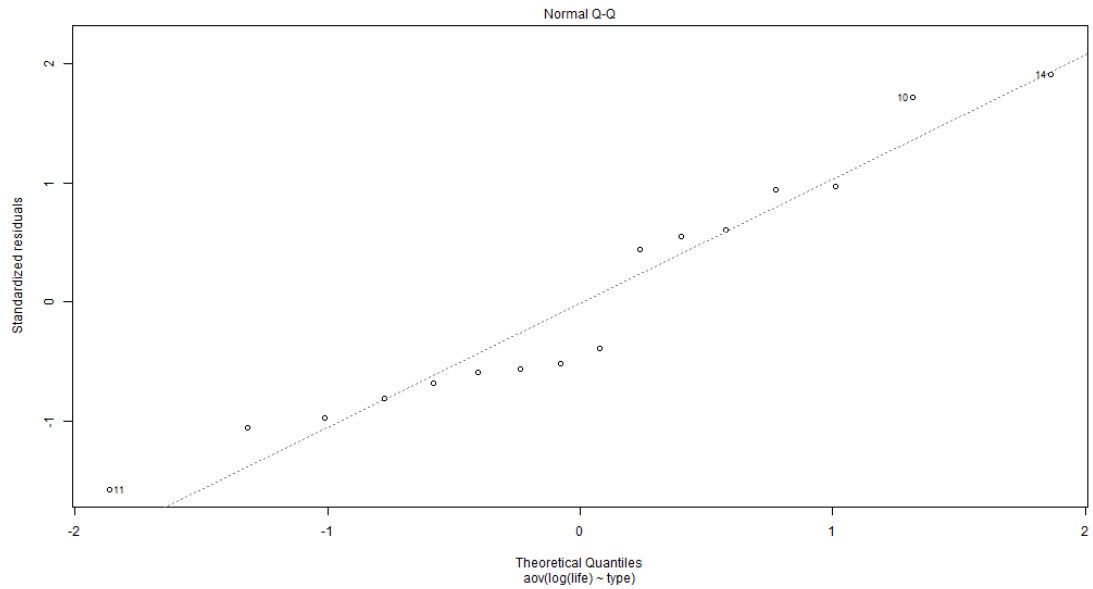
$$\log(Y_{it}) = \mu + \tau_i + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma^2)$$

$$i = \text{AlkName}, \text{AlkGen}, \text{HDName}, \text{HDGen}$$

$$t = 1, \ldots, r_i; \quad r_{AlkName} = r_{AlkGen} = r_{HDName} = r_{HDGen} = 4$$

b)





c)

No. The Residuals vs Fitted plots clearly show an decreasing trend in the residuals.

PROBLEM 5

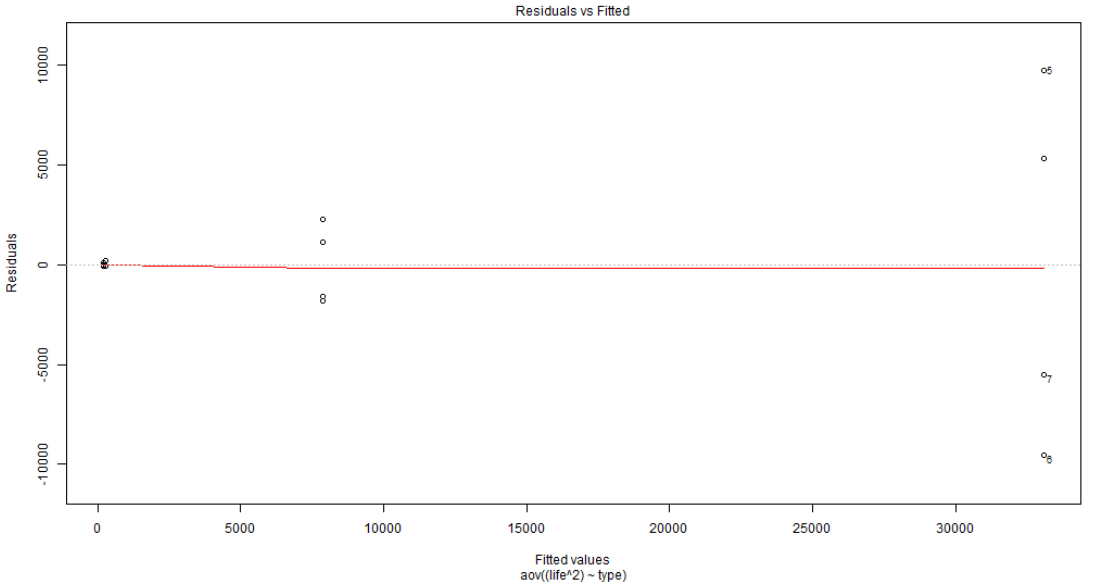
a)

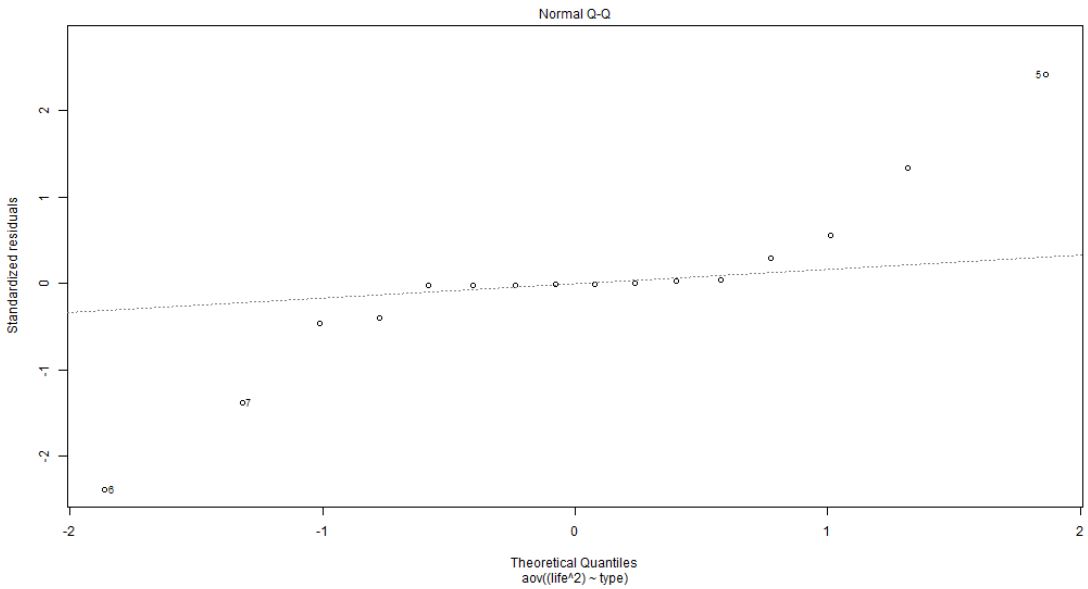
$$Y_{it}^2 = \mu + \tau_i + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma^2)$$

$i = \text{AlkName}, \text{AlkGen}, \text{HDName}, \text{HDGen}$

$t = 1, \ldots, r_i; \quad r_{AlkName} = r_{AlkGen} = r_{HDName} = r_{HDGen} = 4$

b)





c)

No. The Residuals vs Fitted plots clearly show a significantly increasing trend in the residuals.

PROBLEM 6

The square root model clearly has the nicest looking Residuals vs Fitted plots.

As shown below, there are significant differences for every pair other than HD Generic vs HD Name-brand.

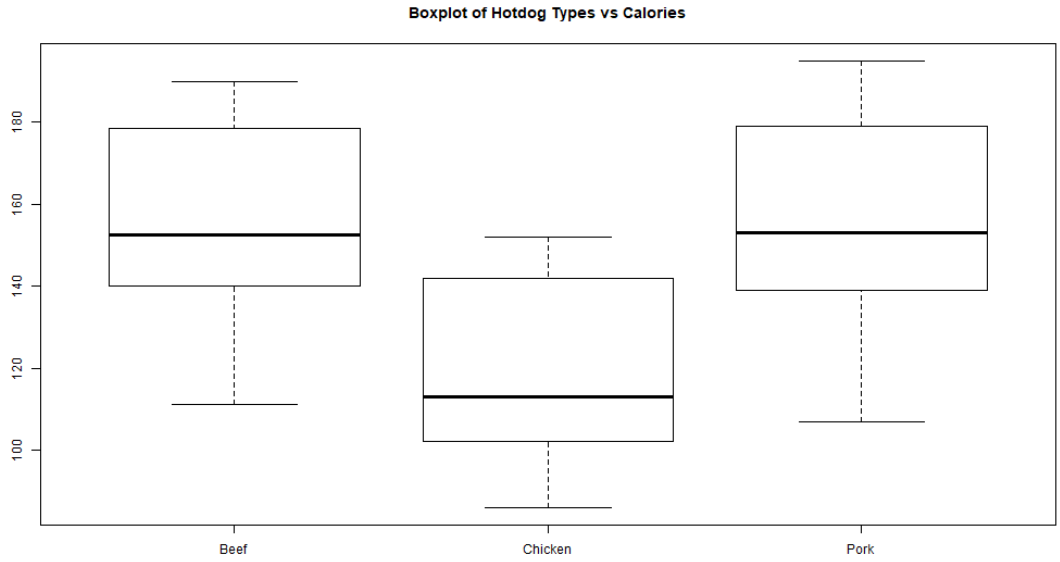
- Alk Gen > Alk Name
- Alk Gen > HD Gen
- Alk Gen > HD Name
- Alk Name > HD Gen
- Alk Name > HD Name
- HD Gen = HD Name

```
> contrast(sqrtLifeLSM, method='pairwise')
 contrast           estimate        SE df t.ratio p.value
 AlkGen - AlkName 4.0374808 0.4427383 12   9.119  <.0001
 AlkGen - HDGen   9.3166850 0.4427383 12  21.043  <.0001
 AlkGen - HDName  9.6541649 0.4427383 12  21.806  <.0001
 AlkName - HDGen  5.2792042 0.4427383 12  11.924  <.0001
 AlkName - HDName 5.6166841 0.4427383 12  12.686  <.0001
 HDGen - HDName   0.3374799 0.4427383 12   0.762  0.8697
```

Listing 1: Pairwise Comparisons for Sqrt Model

PROBLEM 8

Below is a boxplot for the calories vs types of hotdog:



Following this is the pairwise comparison of the types of hotdog vs calories. This comparison shows us that the following differences in types exist:

- Beef has higher calories than chicken
- Beef and pork have no significant calorie difference
- Poor has higher calories than chicken

```
> contrast(hotdogLSM, method='pairwise')
 contrast          estimate       SE df t.ratio p.value
 Beef - Chicken   38.085294 7.738831 51   4.921  <.0001
 Beef - Pork      -1.855882 7.738831 51  -0.240  0.9688
 Chicken - Pork -39.941176 8.046454 51  -4.964  <.0001
```

Listing 2: Pairwise Comparisons for Hotdogs

Code Appendix

```r
################################################################################
#### Setup
################################################################################
## Install and load libraries
# ipak function taken from: https://gist.github.com/stevenworthington/3178163
ipak <- function(pkg) {
  new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
  if (length(new.pkg))
    install.packages(new.pkg, dependencies = TRUE)
  sapply(pkg, require, character.only = TRUE)
}
packages <- c("ggplot2", "reshape2", "gridExtra", "TSA", "astsa", "orcutt",
              "nlme", "fGarch", "vars", "lsmeans")
ipak(packages)

################################################################################
#### Data Input
################################################################################
type<-c("AlkName","AlkName","AlkName","AlkName","AlkGen","AlkGen",
        "AlkGen","AlkGen","HDName","HDName","HDName","HDName",
        "HDGen","HDGen","HDGen","HDGen")

life<-c(100.668, 77.734,79.210,95.063,206.880,153.347,165.980,196.000,
        14.951,18.063,11.111,12.840,15.340,22.090,15.734, 14.440)

batt<-data.frame(type=type, life=life)

################################################################################
#### Problem 1
################################################################################
png("./figures/battery-boxplot.png", width = 1024, height = 576)
  boxplot(life ~ type, main="Boxplot of Battery Data")
dev.off()

################################################################################
#### Problem 2
################################################################################
## Part B
lifeAnova = aov(life~type, data=batt)
summary(lifeAnova)

png("./figures/original-ResidualVsFitted.png", width = 1024, height = 576)
  plot(lifeAnova,1)
dev.off()

png("./figures/original-QQ.png", width = 1024, height = 576)
  plot(lifeAnova,2)
dev.off()

################################################################################
#### Problem 3
################################################################################
## Part B
sqrtLifeAnova = aov(sqrt(life)~type, data=batt)
summary(sqrtLifeAnova)

png("./figures/root-ResidualVsFitted.png", width = 1024, height = 576)
  plot(sqrtLifeAnova,1)
dev.off()

png("./figures/root-QQ.png", width = 1024, height = 576)
  plot(sqrtLifeAnova,2)
dev.off()

```

```
################################################################################
#### Problem 4
################################################################################
## Part B
logLifeAnova = aov(log(life)~type, data=batt)
summary(logLifeAnova)

png("./figures/log-ResidualVsFitted.png", width = 1024, height = 576)
  plot(logLifeAnova,1)
dev.off()

png("./figures/log-QQ.png", width = 1024, height = 576)
  plot(logLifeAnova,2)
dev.off()

################################################################################
#### Problem 5
################################################################################
## Part B
squareLifeAnova = aov((life^2)~type, data=batt)
summary(squareLifeAnova)

png("./figures/square-ResidualVsFitted.png", width = 1024, height = 576)
  plot(squareLifeAnova,1)
dev.off()

png("./figures/square-QQ.png", width = 1024, height = 576)
  plot(squareLifeAnova,2)
dev.off()

################################################################################
#### Problem 7
################################################################################
sqrtLifeLSM = lsmeans(sqrtLifeAnova, ~ type)
contrast(sqrtLifeLSM, method='pairwise')

################################################################################
#### Problem 8
################################################################################
hotdog=read.table("hotdogs.txt",header=TRUE)

png("./figures/hotdog-boxplot.png", width = 1024, height = 576)
  boxplot(hotdog$Calories~hotdog$Type, main="Boxplot of Hotdog Types vs Calories")
dev.off()

hotdogModel = aov(Calories ~ Type, data=hotdog)
summary(hotdogModel)
hotdogLSM = lsmeans(hotdogModel, ~Type)
contrast(hotdogLSM, method='pairwise')
```