Homework3

Aaron Marshall

2023-10-18

Part A

```
library(tidyverse)
```

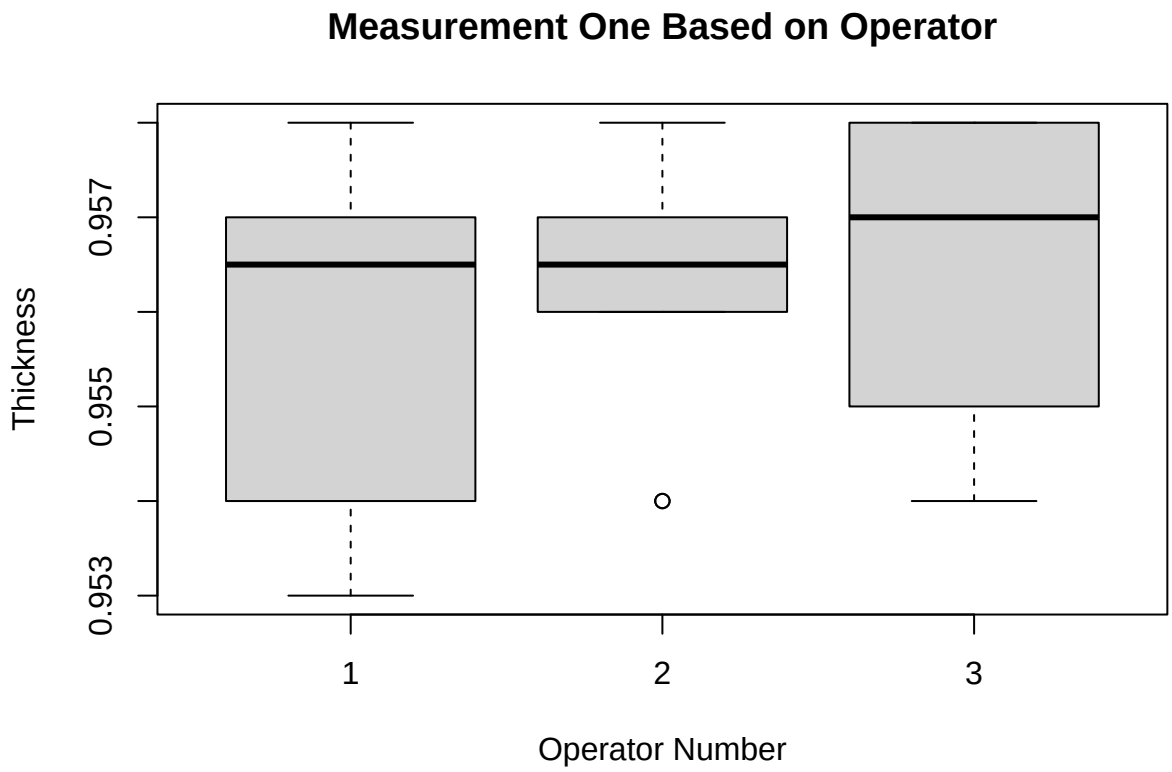
```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.2     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.2     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
urlA = "https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/ThicknessGauge.dat"
dataA<-read.table(urlA, header=F, skip=0, fill=T, stringsAsFactors = F)
dataAFixed <- dataA[-c(1:2),]
dataA2 <- dataAFixed %>% pivot_longer(cols = c(V2,V4,V6), names_to = "Operator", values_to = "Measureme
dataA3 <- dataAFixed %>% pivot_longer(cols = c(V3,V5,V7), names_to = "Operator2", values_to = "Measurem
dataA4 <- cbind(dataA2,dataA3)
dataA5 <- dataA4[,c(1,5,6,12)]
colnames(dataA5)= c("Part", "Operator", "Measurement One", "Measurement Two")
for (i in 1:nrow(dataA5)) {
  if(i%%3==1){
    dataA5[i,2]=1
  } else{
    if(i%%3==2){
      dataA5[i,2]=2
    } else{
      dataA5[i,2]=3
    }
  }
}
dataA5$Part <- as.factor(dataA5$Part)
dataA5$Operator <- as.factor(dataA5$Operator)
summary(dataA5)
```

```
##       Part    Operator Measurement One  Measurement Two
##  1      : 3   1:10     Min.   :0.9530   Min.   :0.952
```

```
##  10      : 3   2:10     1st Qu.:0.9553   1st Qu.:0.955
##  2       : 3   3:10     Median :0.9570   Median :0.956
##  3       : 3            Mean   :0.9563   Mean   :0.956
##  4       : 3            3rd Qu.:0.9570   3rd Qu.:0.957
##  5       : 3            Max.   :0.9580   Max.   :0.958
##  (Other):12
```

```
plot(dataA5$Operator, dataA5$`Measurement One`,
     main = "Measurement One Based on Operator", xlab = "Operator Number",
     ylab = "Thickness")
```

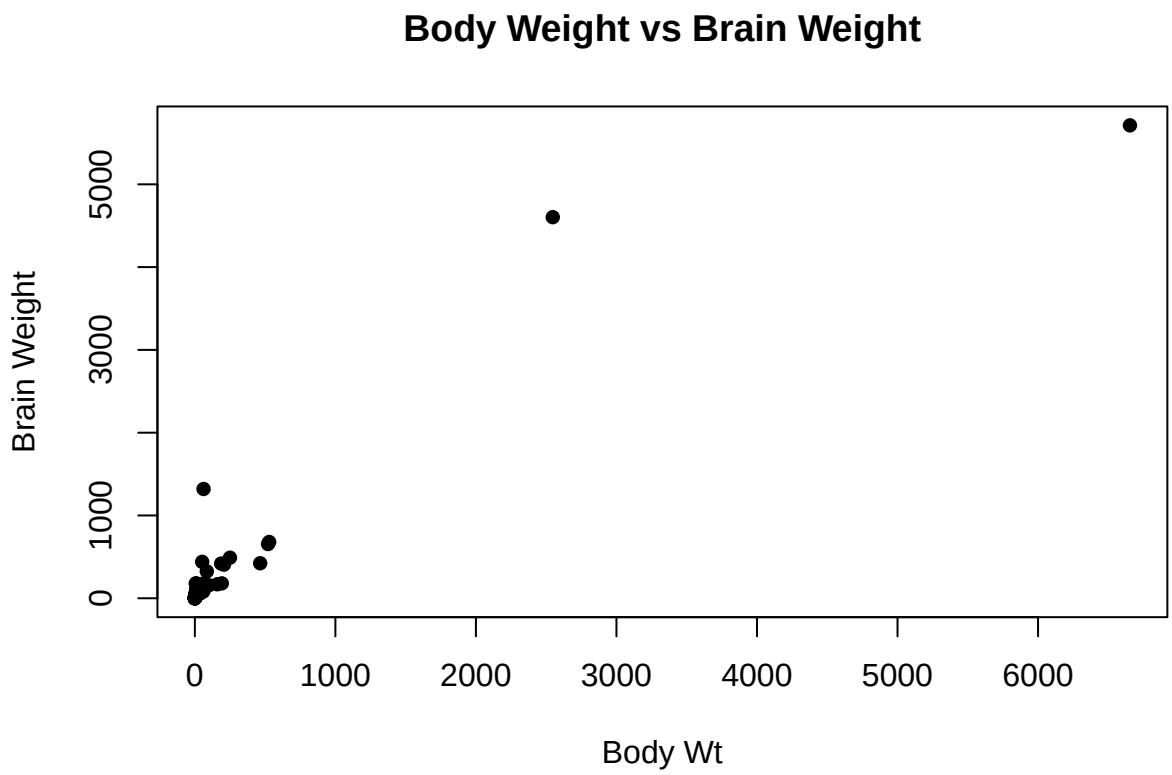


Part B

```
urlB = "https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/BrainandBodyWeight.dat"
dataB<-read.table(urlB, header=T, skip=0, fill=T, stringsAsFactors = F)
dataBFixed <- dataB[,1:6]
FirstBody <- dataB[,1:2]
SecondBody <- dataB[,3:4]
ThirdBody <- dataB[,5:6]
colnames(FirstBody) = c("Body Wt", "Brain Wt")
colnames(SecondBody) = c("Body Wt", "Brain Wt")
colnames(ThirdBody) = c("Body Wt", "Brain Wt")
dataB2 <- rbind(FirstBody,SecondBody,ThirdBody)
dataB3 <- dataB2[-63,]
summary(dataB3)
```

```
##     Body Wt            Brain Wt
##  Min.   :   0.005   Min.   :   0.10
##  1st Qu.:   0.600   1st Qu.:   4.25
##  Median :   3.342   Median :  17.25
##  Mean   : 198.790   Mean   : 283.13
##  3rd Qu.:  48.202   3rd Qu.: 166.00
##  Max.   :6654.000   Max.   :5712.00
```

```
plot(dataB3$`Body Wt`,dataB3$`Brain Wt`, main = "Body Weight vs Brain Weight",
     xlab = "Body Wt", ylab = "Brain Weight", pch = 16)
```



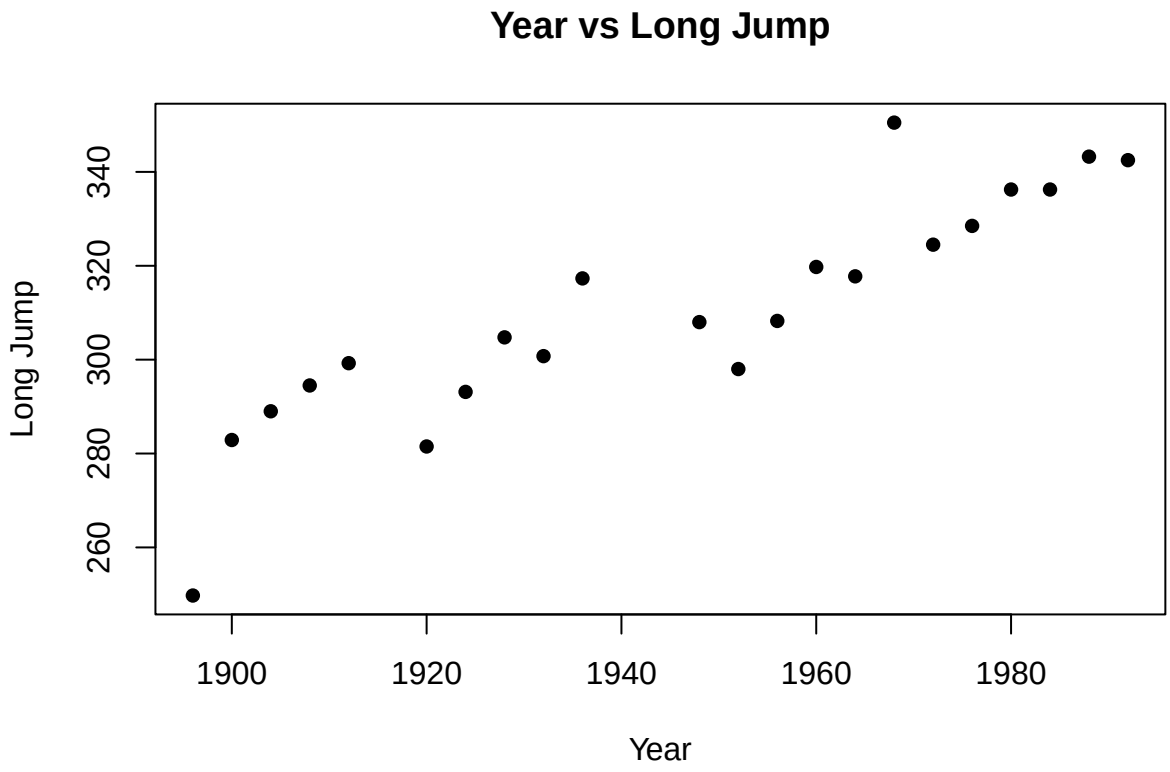
Part C

```
urlC = "https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/LongJumpData.dat"
dataC<-read.table(urlC, header=T, skip=0, fill=T, stringsAsFactors = F)
dataCFixed <- dataC[,1:8]
FirstJump <- dataCFixed[,1:2]
SecondJump <- dataCFixed[,3:4]
ThirdJump <- dataCFixed[,5:6]
FourthJump <- dataCFixed[,7:8]
colnames(FirstJump) = c("Year", "Long Jump")
colnames(SecondJump) = c("Year", "Long Jump")
colnames(ThirdJump) = c("Year", "Long Jump")
colnames(FourthJump) = c("Year", "Long Jump")
dataC2 <- rbind(FirstJump,SecondJump,ThirdJump,FourthJump)
dataC3 <- dataC2[1:22,]
```

```
dataC3$Year <- dataC3$Year+1900
summary(dataC3)
```

```
##       Year        Long Jump
##  Min.   :1896   Min.   :249.8
##  1st Qu.:1921   1st Qu.:295.4
##  Median :1950   Median :308.1
##  Mean   :1945   Mean   :310.3
##  3rd Qu.:1971   3rd Qu.:327.5
##  Max.   :1992   Max.   :350.5
```

```
plot(dataC3$`Year`,dataC3$`Long Jump`, main = "Year vs Long Jump",
     xlab = "Year", ylab = "Long Jump", pch = 16)
```



Part D

```
urlD = "https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/tomato.dat"
dataD<-read.table(urlD, header=F, skip=0, fill=T, stringsAsFactors = F,
                  comment.char ="")
dataDFixed <- dataD[3:4,1:4]
dataD2 <- dataDFixed %>% pivot_longer(cols = V2:V4,
                                      names_to = "Plant Density",
                                      values_to = "Yield")
dataD3 <- dataD2 %>%
          separate(col = "Yield",
```

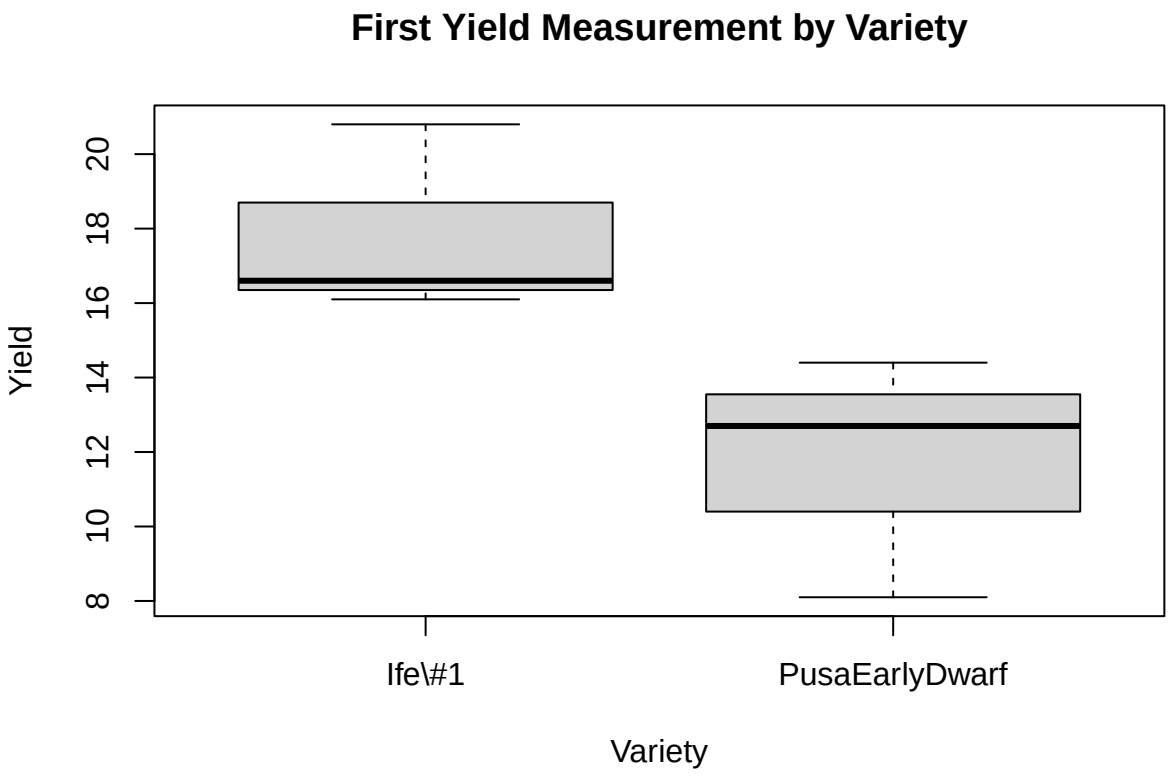
```
                  into = c("Measurement One", "Measurement Two", "Measurement Three"),
                          sep = ",")
```

```
## Warning: Expected 3 pieces. Additional pieces discarded in 1 rows [4].
```

```
for (i in 1:nrow(dataD3)) {
  if(i%%3==1){
    dataD3[i,2]="10000"
  } else{
    if(i%%3==2){
      dataD3[i,2]="20000"
    } else{
      dataD3[i,2]="30000"
    }
  }
}
colnames(dataD3) = c("Variety","Plant Density","Measurement One","Measurement Two", "Measurement Three")
dataD3$Variety <- as.factor(dataD3$Variety)
dataD3$`Plant Density` <- as.factor(dataD3$`Plant Density`)
dataD3$`Measurement One` <- as.double(dataD3$`Measurement One`)
dataD3$`Measurement Two` <- as.double(dataD3$`Measurement Two`)
dataD3$`Measurement Three` <- as.double(dataD3$`Measurement Three`)
summary(dataD3)
```

```
##           Variety   Plant Density Measurement One Measurement Two
##  Ife\\#1       :3   10000:2       Min.   : 8.10   Min.   : 8.60
##  PusaEarlyDwarf:3   20000:2       1st Qu.:13.12   1st Qu.:14.10
##                     30000:2       Median :15.25   Median :15.35
##                                   Mean   :14.78   Mean   :15.03
##                                   3rd Qu.:16.48   3rd Qu.:17.35
##                                   Max.   :20.80   Max.   :19.20
##  Measurement Three
##  Min.   :10.10
##  1st Qu.:12.05
##  Median :15.60
##  Mean   :15.38
##  3rd Qu.:18.25
##  Max.   :21.00
```

```
plot(dataD3$Variety,dataD3$`Measurement One`,
     main = "First Yield Measurement by Variety", xlab = "Variety",
     ylab = "Yield")
```



Part E

```
urlE = "https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/LarvaeControl.dat"
dataE<-read.table(urlE, header=F, skip=0, fill=T, stringsAsFactors = F)
dataEFixed <- dataE[4:11,]
dataEFixed$V2 <- as.double(dataEFixed$V2)
dataEFixed$V3 <- as.double(dataEFixed$V3)
dataE2 <- dataEFixed %>% pivot_longer(cols = c(V2:V11),
                                      names_to = "Treatment",
                                      values_to = "Count")
Age <- c(rep(c(1,1,1,1,1,2,2,2,2,2),8))
dataE3 <- cbind(dataE2, Age)
colnames(dataE3) <- c("Block", "Treatment", "Count", "Age")
for (i in 1:nrow(dataE3)) {
  if(i%%5==1){
    dataE3[i,2]=1
  } else{
    if(i%%5==2){
      dataE3[i,2]=2
    } else{
      if(i%%5==3){
        dataE3[i,2]=3
      } else{
        if(i%%5==4){
          dataE3[i,2]=4
        } else{
```

```
          dataE3[i,2]=5
        }
      }
    }
  }
}
dataE3$Block <- as.factor(dataE3$Block)
dataE3$Treatment <- as.factor(dataE3$Treatment)
summary(dataE3)
```

```
##      Block    Treatment     Count            Age
##  1      :10   1:16      Min.   : 0.00   Min.   :1.0
##  2      :10   2:16      1st Qu.: 2.75   1st Qu.:1.0
##  3      :10   3:16      Median : 5.50   Median :1.5
##  4      :10   4:16      Mean   :10.50   Mean   :1.5
##  5      :10   5:16      3rd Qu.:13.00   3rd Qu.:2.0
##  6      :10             Max.   :61.00   Max.   :2.0
##  (Other):20
```

```
plot(dataE3$Treatment,dataE3$Count,main = "Larvae Counts in Each Treatment",
     xlab = "Treatment", ylab = "Count")
```