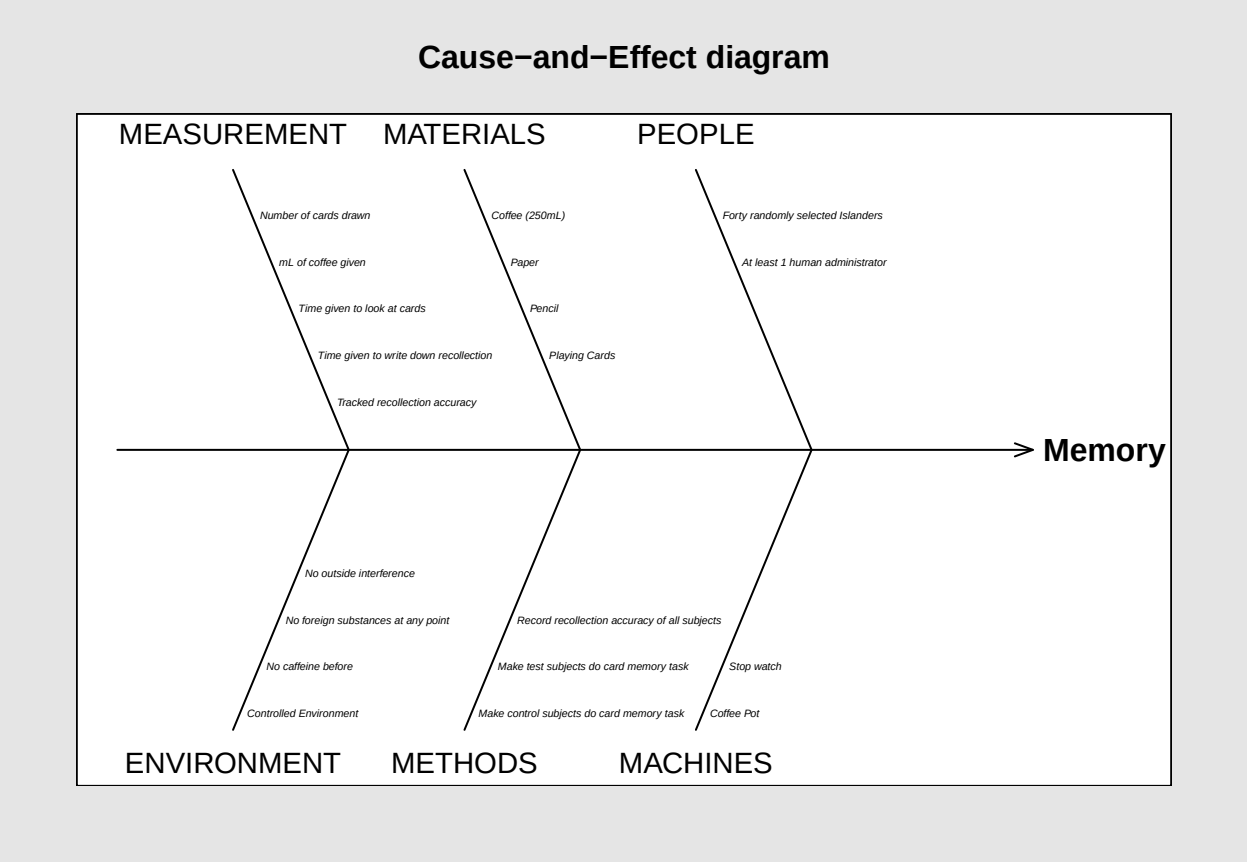
Connor Hennen

804318091

Discussion 1B

1a

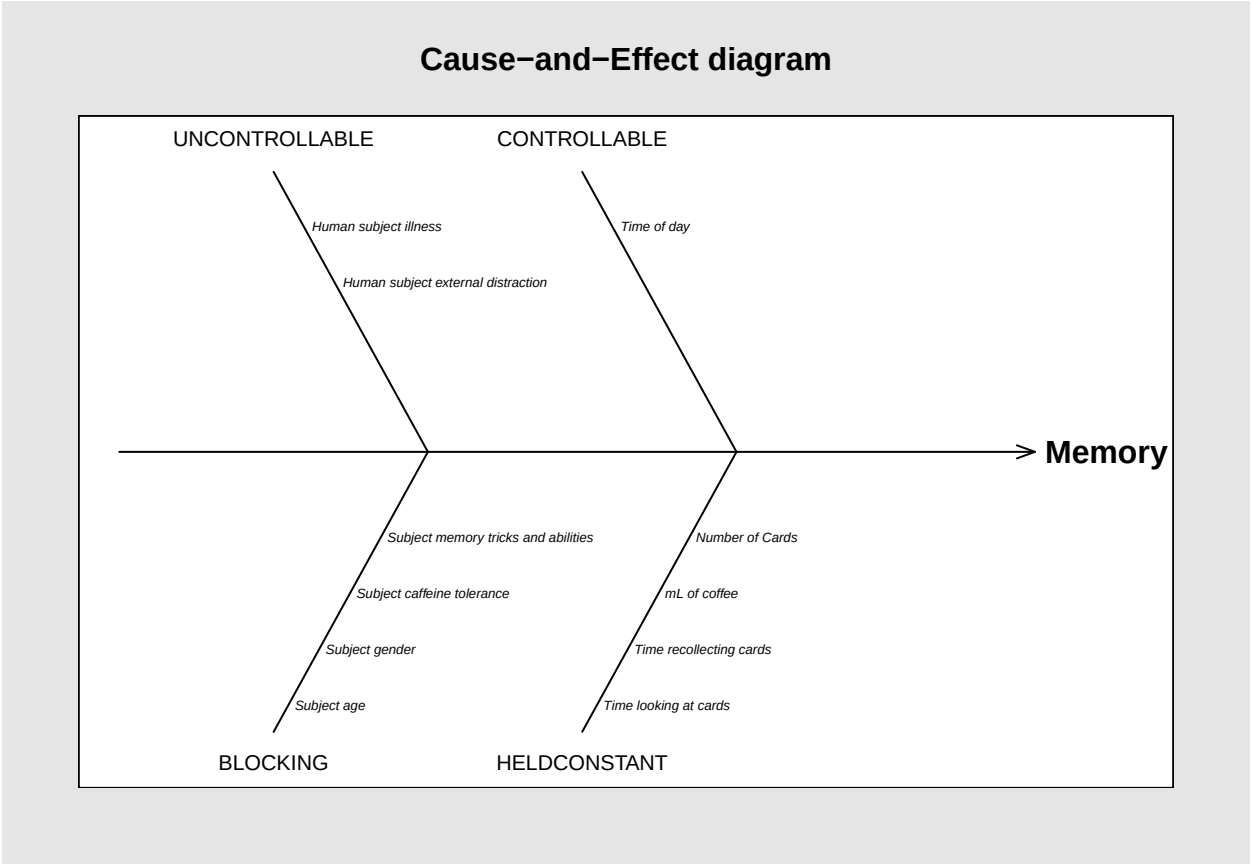
```
library(qcc)
cause.and.effect(
  cause=list(
    MEASUREMENT = c("Number of cards drawn","mL of coffee given",
                    "Time given to look at cards",
                    "Time given to write down recollection",
                    "Tracked recollection accuracy"),
    MATERIALS = c("Coffee (250mL)","Paper","Pencil","Playing Cards"),
    PEOPLE = c("Forty randomly selected Islanders", "At least 1 human administrator"),
    ENVIRONMENT = c("Controlled Environment","No caffeine before",
                    "No foreign substances at any point","No outside interference"),
    METHODS = c("Make control subjects do card memory task",
                "Make test subjects do card memory task",
                "Record recollection accuracy of all subjects"),
    MACHINES = c("Coffee Pot","Stop watch")
  ),
effect="Memory",cex = c(.9, 0.3, 1))
```




```
## There are 6 values in this test that can be varied or measured.
## These are all the items under measurement.
## All of the items under materials are either consumables or items of static items
## of low complexity. One could perhaps consider pencils machines,
## but they are consumable, very basic, and are controlled manually.
## The items under people are all of the people involved in this test.
## The items under environment make it such that it is controlled without outside interference,
##  which, for this test, is all that is needed.
## The items under methods describe how the order and nature of the tasks that are used for this
## experiment.The items under machines are the only mechanical appartus involved
```

1b)

```
library(qcc)
cause.and.effect(
  cause=list(
    UNCONTROLLABLE = c("Human subject illness","Human subject external distraction"),
    CONTROLLABLE = c("Time of day"),
    BLOCKING = c("Subject age", "Subject gender", "Subject caffeine tolerance",
                 "Subject memory tricks and abilities"),
    HELDCONSTANT = c("Time looking at cards","Time recollecting cards",
                     "mL of coffee", "Number of Cards")
  ),
effect="Memory",cex = c(.7, 0.4, 1))
```




```
## Personal or medical problems are often unforseen, unknown, and therefore very difficult to control.
## We can, however, control at what time the test is given and how long the subjects have to rest
## or relax before the test. Blocking variables, which also will cause variances, but are not directly
## related to the hypothesis are age, gender, caffeine tolerance, and memory tricks and abilities.
## All of these vary accross a population, but are not necessarily related to
## coffee's effect on memory. We choose to hold all of the test variables that do relate
## explicitly to coffee, memory, and our materials constant.
```

1c)

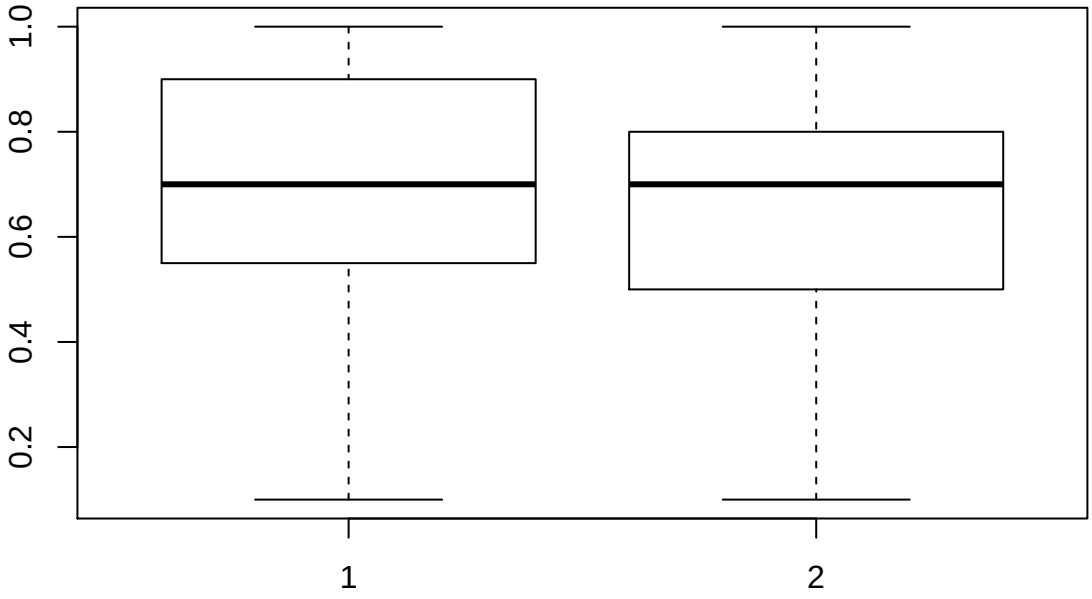
```
##  [1]  2  6  7  8  9 11 13 15 17 18 20 21 22 23 32 33 34 35 38 40
```

```
coffeeScores <- c(0.6, 0.8, 1.0, 0.2, 0.8, 0.7, 0.7, 0.9, 0.9, 0.5, 1, 0.3, 0.9, 0.9, 0.1, 0.5, 0.6, 0.
noCoffeeScores <- c(0.6, 0.6, 0.7, 0.6, 0.2, 1, 0.7, 0.7, 0.4, 0.8, 0.1, 0.8, 0.4, 0.7, 0.7,1, 0.8, 0.7


allScores <- as.double(c(0.6, 0.8, 1.0, 0.2, 0.8, 0.7, 0.7, 0.9, 0.9, 0.5, 1, 0.3, 0.9, 0.9, 0.1, 0.5, 

theTest <- as.character(c(rep("testGroup",20),rep("none",20)))
df <- cbind(allScores, coffeeScores)
df <- as.data.frame(df)
df$allScores <- as.numeric(as.character(df$allScores))

boxplot(coffeeScores, noCoffeeScores)
```




```
print("We can see very similar medians for the two groups but the third and
      first quartiles are better for coffee group. However, probably too
      small to be statistically significant given our sample sizes.")
```

```
## [1] "We can see very similar medians for the two groups but the third and \n      first quartiles ar
```

```
t.test(coffeeScores,noCoffeeScores)
```

```
##
##  Welch Two Sample t-test
##
## data:  coffeeScores and noCoffeeScores
## t = 0.65893, df = 37.927, p-value = 0.5139
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1139853  0.2239853
## sample estimates:
## mean of x mean of y
##     0.685     0.630
```

```
print("With a p-value 0f 0.5139, we cannot confirm the that coffee improves memory.")
```

```
## [1] "With a p-value 0f 0.5139, we cannot confirm the that coffee improves memory."
```

1d)

```
clm <- lm(df$allScores ~ df$coffeeScores)
anova(clm)
```

```
## Analysis of Variance Table
##
## Response: df$allScores
##                  Df  Sum Sq  Mean Sq F value  Pr(>F)
## df$coffeeScores   1 0.30215 0.302153  4.8332 0.03408 *
## Residuals        38 2.37560 0.062516
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
print("Again, we can see via the F-statistic, that coffee cannot definitely improve memory.")
```

```
## [1] "Again, we can see via the F-statistic, that coffee cannot definitely improve memory."
```

2

```
error_ss <- 196.14 - 36.15
factor_ms <- 36.15/3
error_ms <- error_ss/16
fvalue <- factor_ms/error_ms
pvalue <- pf(fvalue,3,16)
df2 <- c(3,16,19)
ss <- c(36.15, error_ss, 196.14)
ms <- c(factor_ms,error_ms,NA)
p <- c(fvalue, NA, NA)
q <- c(pvalue, NA, NA)
table1 <- cbind(df2,ss,ms,p,q)
table1
```

```
##      df2     ss        ms        p         q
## [1,]   3  36.15 12.050000 1.205075 0.6602178
## [2,]  16 159.99  9.999375       NA        NA
## [3,]  19 196.14        NA       NA        NA
```

3

```
HW2_Q3_Data <- read.csv("~/Downloads/HW2 Q3 Data.csv")

sham <- HW2_Q3_Data$Sham
h1 <- HW2_Q3_Data$PEMF.1h.day
h2 <- HW2_Q3_Data$PEMF.2h.day
h4 <- HW2_Q3_Data$PEMF.4h.day
all <- c(sham,h1,h2,h4)
names2 <- c(rep("sham",20),rep("h1",20),rep("h2",20),rep("h4",20))
arm_results <- as.data.frame(cbind(names2, all))
arm_results$names2 <- as.character(arm_results$names2)
arm_results$all <- as.numeric(as.character(arm_results$all))
arm_lm <- lm(arm_results$all ~ arm_results$names2) #summary(arm_lm)
anova(arm_lm)
```

```
## Analysis of Variance Table
##
## Response: arm_results$all
##                     Df  Sum Sq Mean Sq F value Pr(>F)
## arm_results$names2   3  10.044  3.3478  1.2979 0.2813
## Residuals           76 196.030  2.5793
```

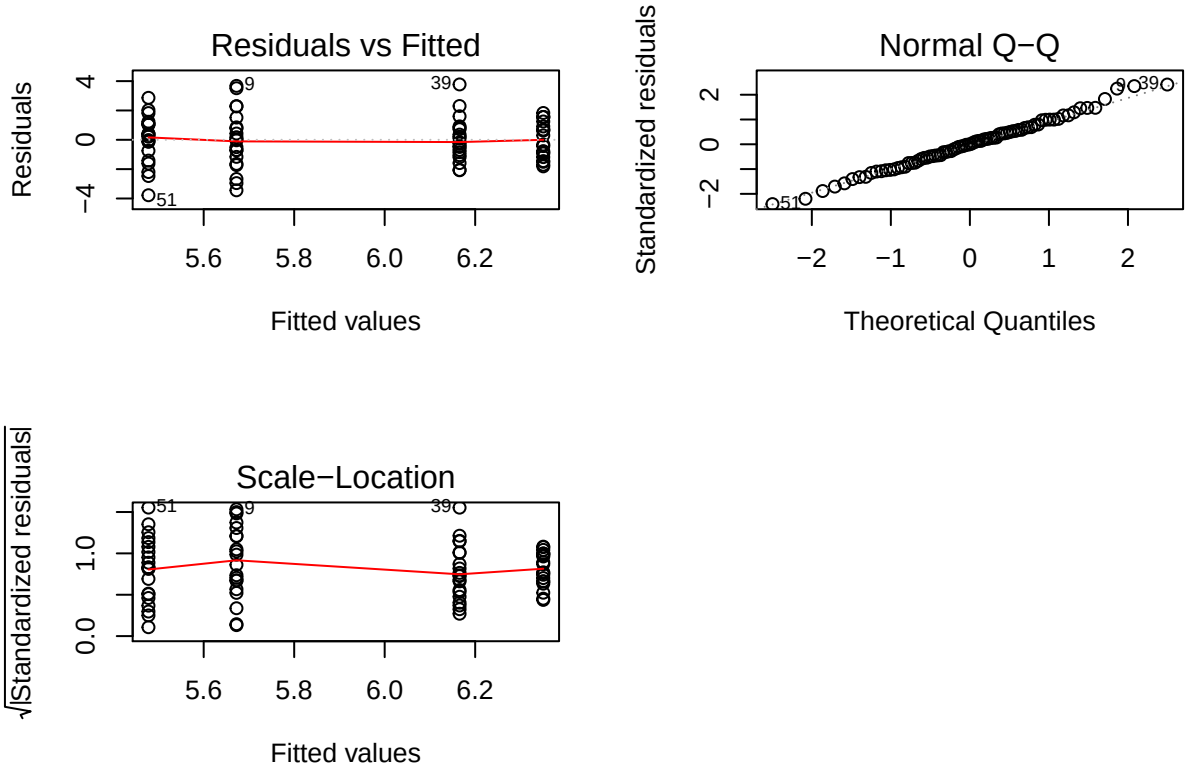
```
print("PEMF usage does not affect BMD loss at a 5% level. p=0.28  which is less than 0.05.")
```

```
## [1] "PEMF usage does not affect BMD loss at a 5% level. p=0.28  which is less than 0.05."
```

3b)

```
par(mfrow = c(2,2))
plot(arm_lm)
```

```
## hat values (leverages) are all = 0.05
##  and there are no factor predictors; no plot no. 5
```




```
print("All of the residual graphs looks pretty good. In the Scale-Location
      graph, there are some sharp nonlinear points but that's sort of no
      surprise given the amount of observations (low amount) we have.")
```

```
## [1] "All of the residual graphs looks pretty good. In the Scale-Location \n      graph, there are so
```

4a

```
roddingLevel <- list(10,15,20,25)

strength1 <- list(1530,1610,1560,1500)
strength2 <- list(1530,1650,1730,1490)
strength3 <- list(1440,1500,1530,1510)
fullData <- as.data.frame(cbind(unlist(roddingLevel),unlist(strength1),unlist(strength2),unlist(strengt
fullData
```

```
##   V1   V2   V3   V4
## 1 10 1530 1530 1440
## 2 15 1610 1650 1500
## 3 20 1560 1730 1530
## 4 25 1500 1490 1510
```

```
level10 <- as.numeric(fullData[1,2:4])
level15 <- as.numeric(fullData[2,2:4])
level20 <- as.numeric(fullData[3,2:4])
level25 <- as.numeric(fullData[4,2:4])

allLevelData <- c(level10,level15,level20,level25)

allLevels<- c(rep("10",3),rep("15",3),rep("20",3),rep("25",3))
levelsAndScores <- cbind(allLevels,allLevelData)
levelsAndScores <- as.data.frame(levelsAndScores)
levelsAndScores$allLevels <- as.character(levelsAndScores$allLevels)
levelsAndScores$allLevelData <- as.numeric(as.character(levelsAndScores$allLevelData))
m1 <- lm(levelsAndScores$allLevelData ~ levelsAndScores$allLevels)
anova(m1)
```

```
## Analysis of Variance Table
##
## Response: levelsAndScores$allLevelData
##                            Df Sum Sq Mean Sq F value Pr(>F)
## levelsAndScores$allLevels  3  28633  9544.4  1.8654 0.2138
## Residuals                  8  40933  5116.7
```

```
print("Anova table tells us that there isn't a statistically significant
      diffence in compressive strength at A = 0.05")
```

```
## [1] "Anova table tells us that there isn't a statistically significant\n      diffence in compressive
```

4b

```
fourB <- anova(m1)
fourB$`F value`[1]
```

```
## [1] 1.865364
```

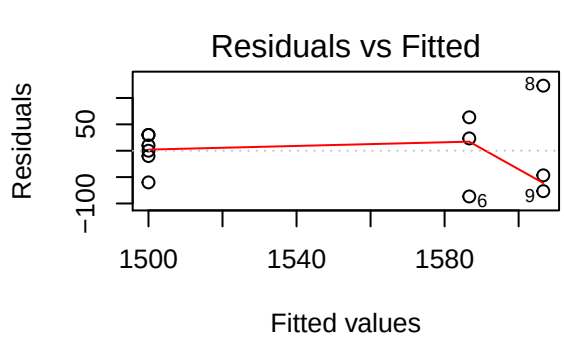
```
cat("The P-value for the F statistic in part (a) is", fourB$`Pr(>F)`[1])
```

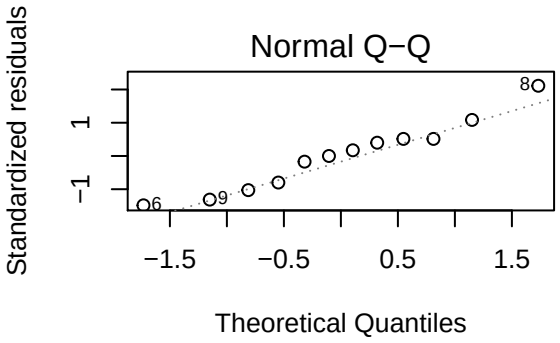
```
## The P-value for the F statistic in part (a) is 0.2137815
```

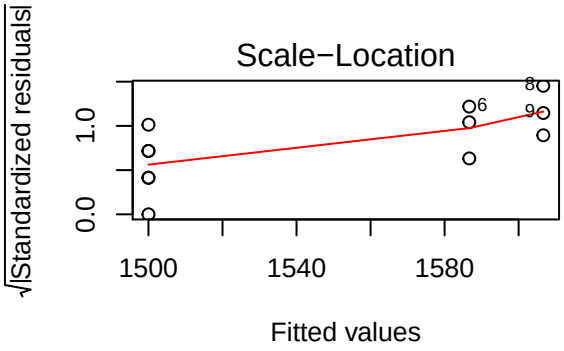
4c

```
par(mfrow = c(2,2))
plot(m1)
```

```
## hat values (leverages) are all = 0.3333333
##  and there are no factor predictors; no plot no. 5
```





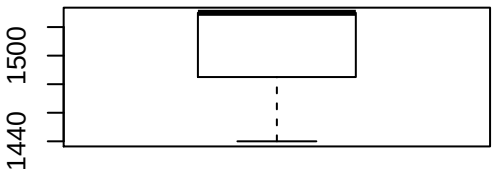


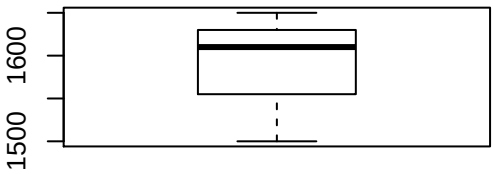

```
print("You can't really draw any conclusions - the residual graphs look
      pretty poor and there are just too few observations")
```

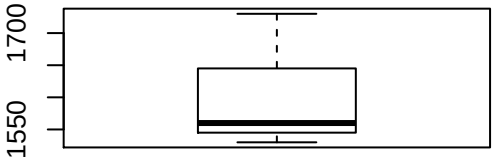
```
## [1] "You can't really draw any conclusions - the residual graphs look \n      pretty poor and there a
```

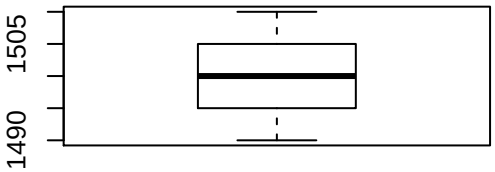
4d

```
par(mfrow=c(2,2))
boxplot(level10)
boxplot(level15)
boxplot(level20)
boxplot(level25)
```


1440
1500


1500
1600


1550
1700


1490
1505