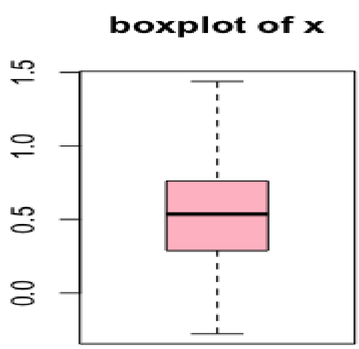
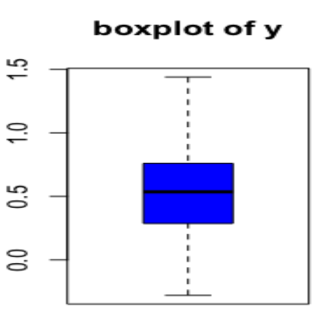
Module 3 Assignment VIVEK PAKALAPATI U49253220

```
#VIVEK PAKALAPATI
#U49253220

rm(list=ls())
install.packages(rio)
install.packages("rio")
library(rio)
library(moments)
attach(randu)
#Last four digits of the U number is 3220
set.seed(3220)
data<-randu + rnorm(400)/7
S= sum(c(3,2,2,0))
data<-randu + rnorm(400)/7
boxplot(data$x, col=c("pink"), main="boxplot of x")
```




```
boxplot(data$x, col=c("blue"), main="boxplot of y")
```



```
boxplot(data$z, col=c("white"), main="box plot of z")
```

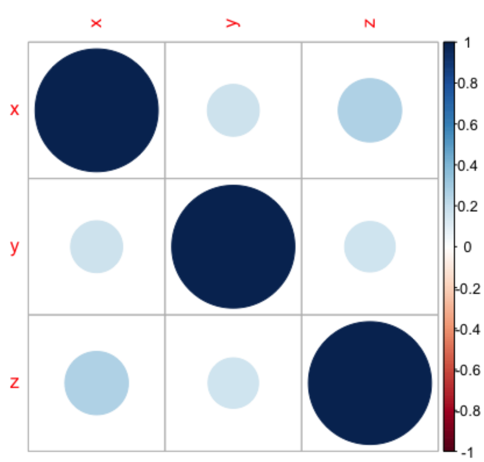
#We can see that all three algorithms have almost similar boxplots with identical values.
#One can understand that the 3 boxplots are similar and range from 0.1 to 0.5.
#As I have used my own U number for generating the box plot I believe everyone will have a unique box plot with unique value. I have also tried with random U numbers and generated box plots and found that the values change and it is unique.

#2

```
set.seed(3220) #last 4 digits of your U-number
before <-c(230.1, 220.9, 212.7, 293, 341.4, 296.9, 192.2, 175.5, 255.2, 293.7)
# Performance after Treatment
after <-c(392.9, 393.2, 345.1, 393, 434, 427.9, 422, 383.9, 392.3, 352.2)
# Create a data frame
mydata <- data.frame(
  group = rep(c("before", "after"), each = 10),
  unit = c(seq(1:10),seq(1:10)),
  score = c(before,  after)
)


install.packages("corrplot")
library(corrplot)

cor_of_3 <-cor(data)
corrplot(cor_of_3)
```



```
t.test(x=data$x, y=data$y)
t.test(x=data$y, y=data$z)
t.test(x=data$x, y=data$z)
```

```
> t.test(x=data$x, y=data$y)

	Welch Two Sample t-test

data:  data$x and data$y
t = 1.7472, df = 797.43, p-value = 0.08099
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.004986326  0.085738796
sample estimates:
mean of x mean of y
0.5284508 0.4880746

> t.test(x=data$y, y=data$z)

	Welch Two Sample t-test

data:  data$y and data$z
t = 0.22268, df = 796.32, p-value = 0.8238
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.03984483  0.05004163
sample estimates:
mean of x mean of y
0.4880746 0.4829762

> t.test(x=data$x, y=data$z)

	Welch Two Sample t-test

data:  data$x and data$z
t = 2.0137, df = 797.7, p-value = 0.04437
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.001147201 0.089802064
sample estimates:
mean of x mean of y
0.5284508 0.4829762
```

```
head(mydata)
```

```
>head(mydata)
  group unit score
1 before    1 230.1
2 before    2 220.9
3 before    3 212.7
4 before    4 293.0
5 before    5 341.4
6 before    6 296.9
```

```
head(mydata)

mydata2 <- mydata[sample(1:nrow(mydata)),]
mydata2 <- mydata2[1:19,]

#pre processing

mydata2 <- mydata[sample(1:nrow(mydata)),]
mydata2 <- mydata[1:nrow(mydata),]

before_mydata2=subset(mydata2,mydata2$group=="before")
before_mydata2=before_mydata2[order(before_mydata2$unit),]
after_mydata2=subset(mydata2,mydata2$group=="after")
after_mydata2=after_mydata2[order(after_mydata2$unit),]


t.test(before_mydata2$score, after_mydata2$score,paired=TRUE)
```

```
t.test(before_mydata2$score, after_mydata2$score,paired=TRUE)

        Paired t-test

data:  before_mydata2$score and after_mydata2$score
t = -8.57, df = 9, p-value = 1.272e-05
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 -180.102 -104.878
sample estimates:
mean difference
       -142.49
```

#So we can conclude that a significant difference exists as the P value is 0.000012 which is less than 0.5. So we can reject the null hypothesis in this case.

#Now there is proper evidence to tell the mean jump height has changed before and after the training program.