Project Exploratory Data Analysis

Krista Dodd

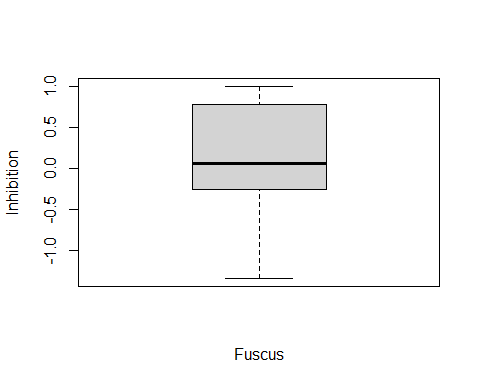
3/3/2021

knitr::opts\_chunk$set(echo = TRUE)

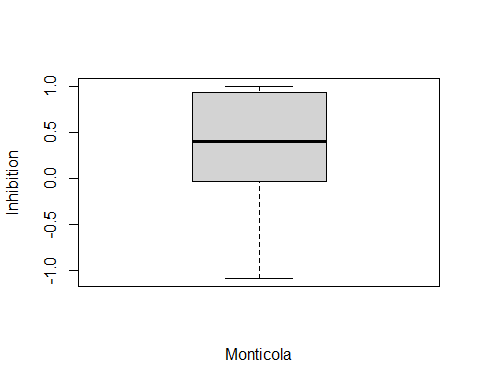
rm(list=ls())  
FuscusMonticola=read.csv("FuscusMonticola.csv",header=TRUE)  
names(FuscusMonticola)

## [1] "Fuscus.Isolate" "Fuscus.Inhibition" "Monticola.Isolate"   
## [4] "Monticola.Inhibition"

boxplot(FuscusMonticola$Fuscus.Inhibition,xlab="Fuscus",ylab="Inhibition")



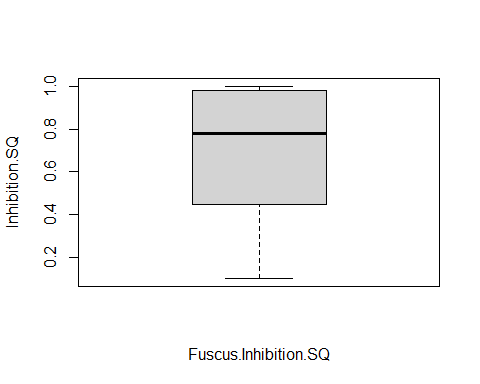
boxplot(FuscusMonticola$Monticola.Inhibition,xlab="Monticola",ylab="Inhibition")



FuscusMonticola$SQ.Fuscus.Inhibition=sqrt(FuscusMonticola$Fuscus.Inhibition)

## Warning in sqrt(FuscusMonticola$Fuscus.Inhibition): NaNs produced

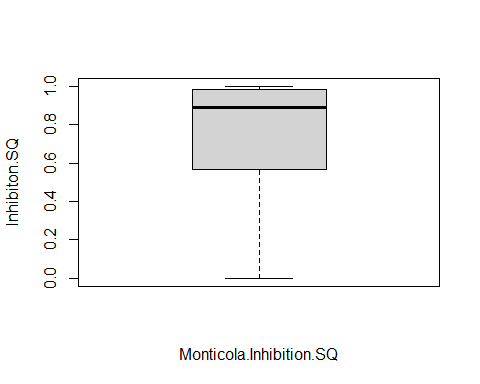
boxplot(FuscusMonticola$SQ.Fuscus.Inhibition,xlab="Fuscus.Inhibition.SQ",ylab="Inhibition.SQ")



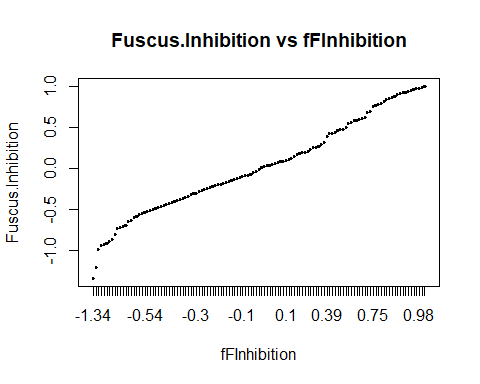
FuscusMonticola$SQ.Monticola.Inhibition=sqrt(FuscusMonticola$Monticola.Inhibition)

## Warning in sqrt(FuscusMonticola$Monticola.Inhibition): NaNs produced

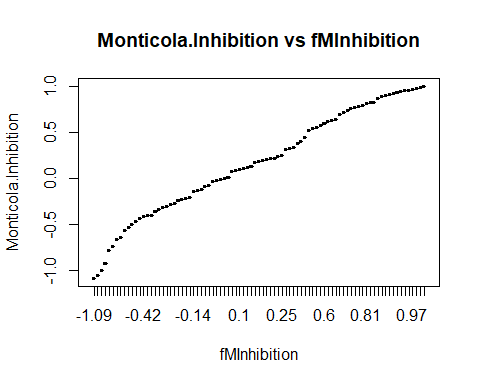
boxplot(FuscusMonticola$SQ.Monticola.Inhibition,xlab="Monticola.Inhibition.SQ",ylab="Inhibiton.SQ")



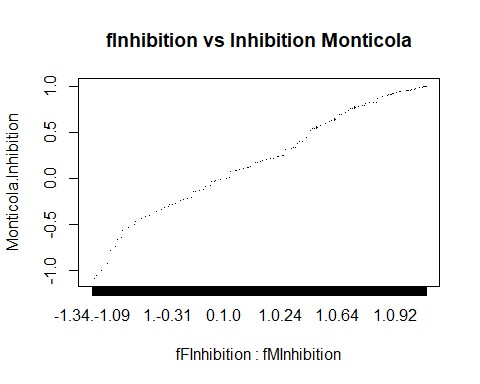
#The datset is unable to be square root transformed because of negative values, therefore it did not help this dataset  
FuscusMonticola$fFInhibition=factor(FuscusMonticola$Fuscus.Inhibition)   
FuscusMonticola$fMInhibition=factor(FuscusMonticola$Monticola.Inhibition)  
boxplot(Fuscus.Inhibition ~ fFInhibition,data=FuscusMonticola,main="Fuscus.Inhibition vs fFInhibition")



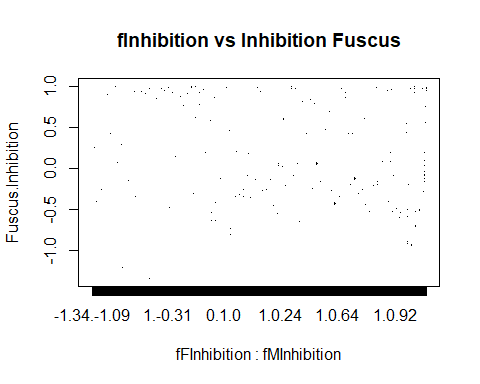
boxplot(Monticola.Inhibition ~ fMInhibition,data=FuscusMonticola,main="Monticola.Inhibition vs fMInhibition")



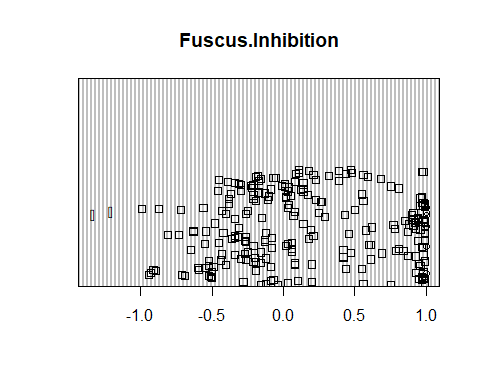
boxplot(Monticola.Inhibition ~ fFInhibition\*fMInhibition,data=FuscusMonticola,main="fInhibition vs Inhibition Monticola")



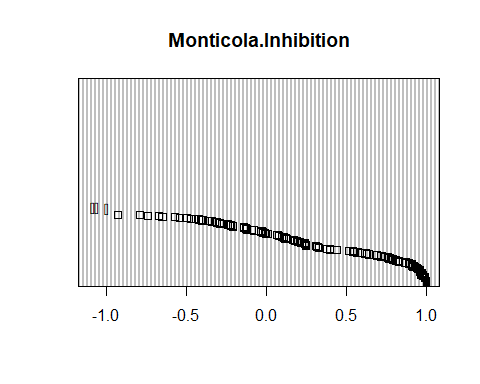
boxplot(Fuscus.Inhibition ~ fFInhibition\*fMInhibition,data=FuscusMonticola,main="fInhibition vs Inhibition Fuscus")



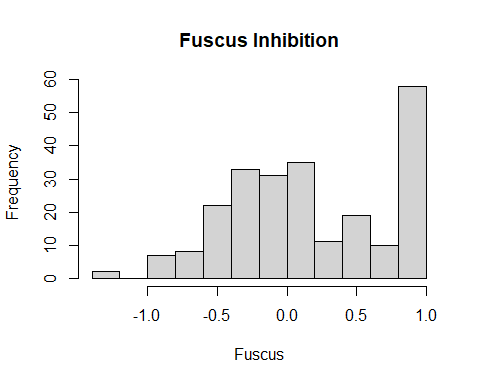
dotchart(FuscusMonticola$Fuscus.Inhibition,pch=as.numeric(FuscusMonticola$Fuscus.Inhibition),main="Fuscus.Inhibition")



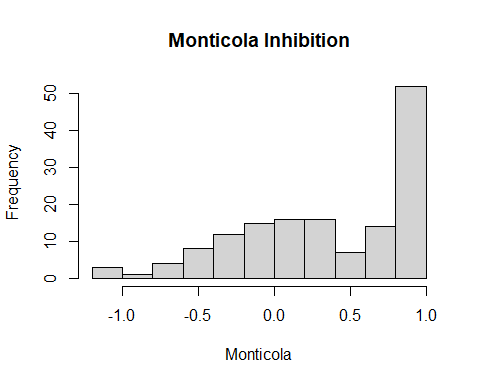
dotchart(FuscusMonticola$Monticola.Inhibition,pch=as.numeric(FuscusMonticola$Monticola.Inhibition),main="Monticola.Inhibition")



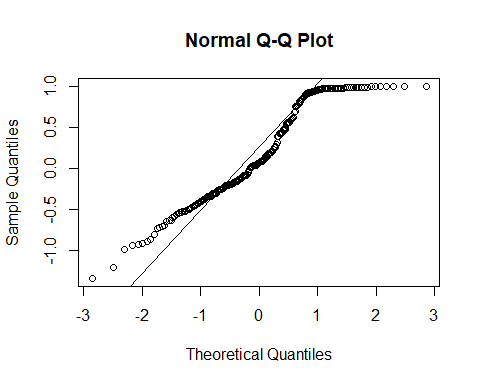
hist(FuscusMonticola$Fuscus.Inhibition,xlab="Fuscus",main="Fuscus Inhibition")



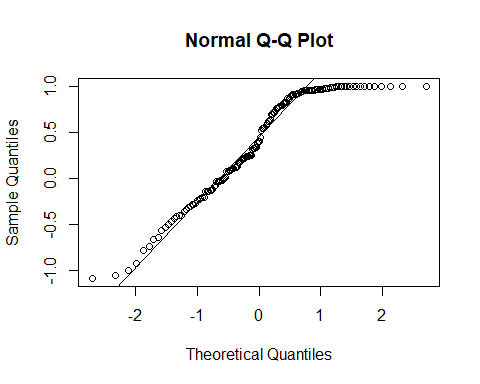
hist(FuscusMonticola$Monticola.Inhibition,xlab="Monticola",main="Monticola Inhibition")



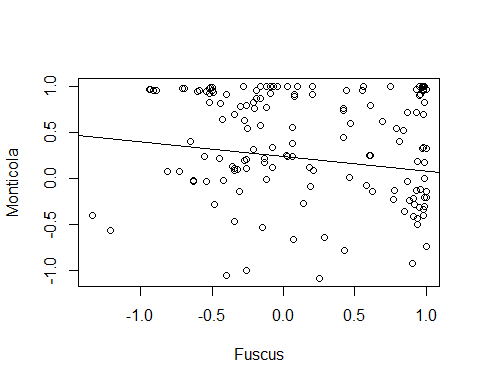
#These datasets look somewhat normally distributed, but not all the way because there were a lot of 1 numbers which pulled the distribution toward that  
qqnorm(FuscusMonticola$Fuscus.Inhibition)  
qqline(FuscusMonticola$Fuscus.Inhibition)



qqnorm(FuscusMonticola$Monticola.Inhibition)  
qqline(FuscusMonticola$Monticola.Inhibition)



#The qq plots do not look very normally distributed either, however the Monticola is better than the Fuscus  
plot(FuscusMonticola$Fuscus.Inhibition,FuscusMonticola$Monticola.Inhibition,xlab="Fuscus",ylab="Monticola")  
mymodel=lm(Fuscus.Inhibition~Monticola.Inhibition,data=FuscusMonticola)  
abline(mymodel)



#There doesn't seem to be a strong relationship here with the scatter plot.   
#My undergrad professor is reorganizing the datasets to make it easier for me to run tests on in the future. I will be uploading that new dataset on Friday.