Assignmen 2

Wenhao Li 250959325

09/02/2020

#Question 1

#(a)

v1<-c(7,21)  
d1<-data.frame(x=1:2,y=c("a","b"))

#(b)

write(v1,file = "data1")  
  
  
write.table(v1,file = "data3")  
write.table(d1,file = "data4")

#(c)

v2<-scan(file = "data1")  
v3<-read.table(file = "data3")  
d2<-read.table(file = "data4")

#(d)

v1

## [1] 7 21

v2

## [1] 7 21

v3

## x  
## 1 7  
## 2 21

all.equal(v1,v2)

## [1] TRUE

all.equal(v1,v3)

## [1] "Modes: numeric, list"   
## [2] "Lengths: 2, 1"   
## [3] "names for current but not for target"   
## [4] "Attributes: < target is NULL, current is list >"  
## [5] "target is numeric, current is data.frame"

d1

## x y  
## 1 1 a  
## 2 2 b

d2

## x y  
## 1 1 a  
## 2 2 b

all.equal(d1,d2,check.attributes=FALSE)

## [1] TRUE

When use scan and write function, they keep the result same for vector object. However, write function can’t export data frame object. When use write.table and read.table, they keep the result same for data.frame object. However, they make the vector object changed. Vector object changed into data.frame object.

#Question 2

data1<-2^(1:15)  
data2<-(1:15)^3  
index<-(1:15)[data1>data2]  
elements<-data1[index]  
index

## [1] 1 10 11 12 13 14 15

elements

## [1] 2 1024 2048 4096 8192 16384 32768

#Question 3

#Similarity: Both dump and save can save objects  
#Difference:   
# "Dump" function takes a vector of names of R objects and produces text representations of the objects on a file or connection. (1)  
# "Save" function writes an external representation of R objects to the speified file.  
  
#creat two objects  
o1<-(1:5)  
o2<-(6:10)  
#dump and source  
dump(list=c("o1","o2"),file = "o3")  
#rename  
new\_o1<-o1  
new\_o2<-o2  
  
source(file = "o3")  
#compare   
all.equal(o1,new\_o1)

## [1] TRUE

all.equal(o2,new\_o2)

## [1] TRUE

#save and load  
o4<-(1:5)  
o5<-(6:10)  
save(list = c("o4","o5"),file = "o6")  
new\_o4<-o4  
new\_o5<-o5  
load(file="o6")  
all.equal(o4,new\_o4)

## [1] TRUE

all.equal(o5,new\_o5)

## [1] TRUE

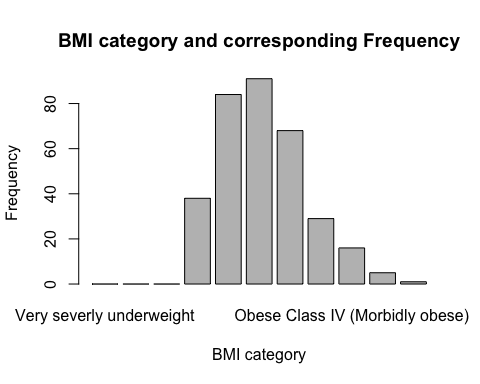
Dump and source get the same R objects back. Save and load get the same R objects back.

#Question 4

library(MASS)  
data(Pima.te)  
BMI.category<-c("Very severly underweight","Severely underweight","Underweight","Normal (healthy weight)","Overweight","Obese Class I (Moderately)","Obese Class II (Severely obese)","Obese Class III (Very severely obese)","Obese Class IV (Morbidly obese)","Obese Class V (Super obese)","Obese Class VI (Hyper obese)")  
BMI.points<-c(0,15,16,18.5,25,30,35,40,45,50,60,Inf)  
BMI\_data<-cut(Pima.te$bmi,breaks = BMI.points,labels = BMI.category,ordered\_result = TRUE )  
table(BMI\_data)

## BMI\_data  
## Very severly underweight Severely underweight   
## 0 0   
## Underweight Normal (healthy weight)   
## 0 38   
## Overweight Obese Class I (Moderately)   
## 84 91   
## Obese Class II (Severely obese) Obese Class III (Very severely obese)   
## 68 29   
## Obese Class IV (Morbidly obese) Obese Class V (Super obese)   
## 16 5   
## Obese Class VI (Hyper obese)   
## 1

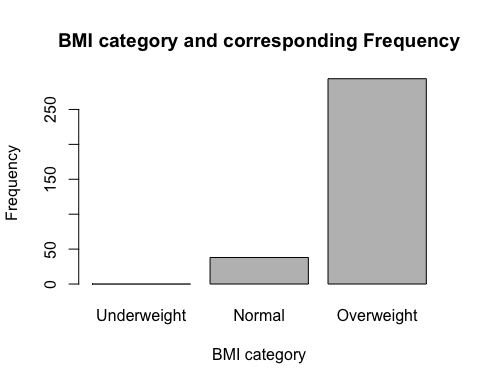
barplot(table(BMI\_data),main = "BMI category and corresponding Frequency",xlab = "BMI category",ylab = "Frequency")



#Cut again  
BMI.category.new<-c("Underweight","Normal","Overweight")  
BMI.points.new<-c(0,18.5,25,Inf)  
BMI\_data.new<-cut(Pima.te$bmi,breaks = BMI.points.new,labels = BMI.category.new,ordered\_result = TRUE)  
table(BMI\_data.new)

## BMI\_data.new  
## Underweight Normal Overweight   
## 0 38 294

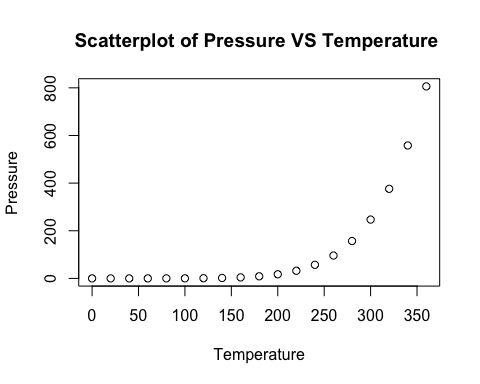
barplot(table(BMI\_data.new),main = "BMI category and corresponding Frequency",xlab = "BMI category",ylab = "Frequency")



#Question 5

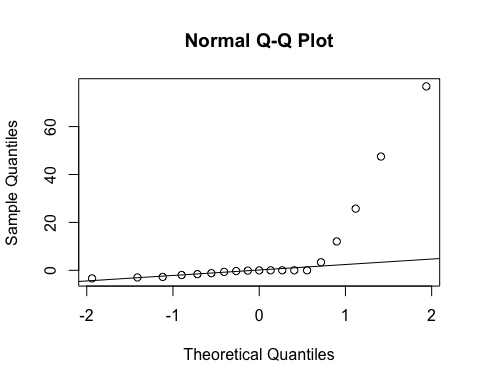
#(a)

data(pressure)  
plot(x=pressure$temperature,y=pressure$pressure,main="Scatterplot of Pressure VS Temperature ",xlab="Temperature",ylab="Pressure")

 The variables(temperature and pressure) related nonlinear.

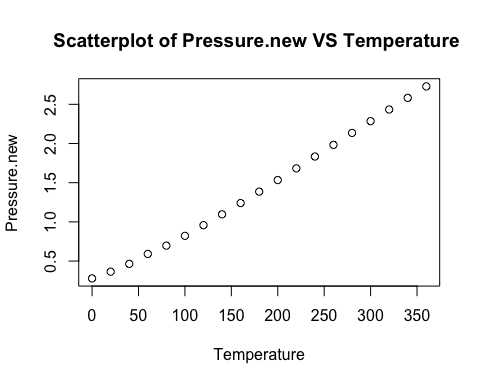
#(b)

residuals<-with(pressure,pressure-(0.168+0.007\*temperature)^(20/3))  
qqnorm(residuals)  
qqline(residuals)

 The residuals are not normally distributed. They follow a skewed distribution.

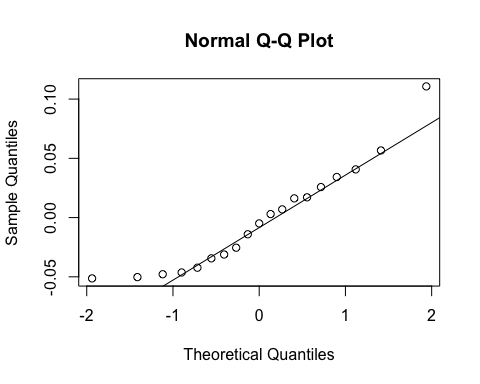
#(c)

pressure.new<-pressure$pressure^(3/20)  
plot(x=pressure$temperature,y=pressure.new,main="Scatterplot of Pressure.new VS Temperature ",xlab="Temperature",ylab="Pressure.new")

 The relationship between transformed pressure values and temperature values is a linear relationship.

#(d)

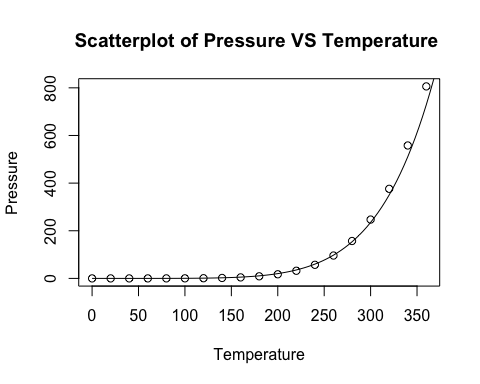
residuals.new<-with(pressure,pressure.new-(0.168+0.007\*temperature))  
qqnorm(residuals.new)  
qqline(residuals.new)

 The residuals are not normally distributed. They follow a skewed distribution.

#Question6

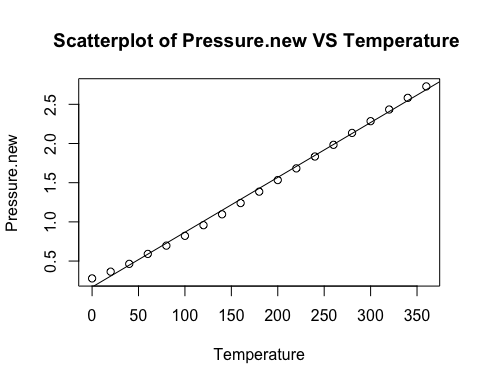
#(a)

plot(x=pressure$temperature,y=pressure$pressure,main="Scatterplot of Pressure VS Temperature ",xlab="Temperature",ylab="Pressure")  
curve((0.168+0.007\*x)^(20/3),from = 0,to=400,add=TRUE)



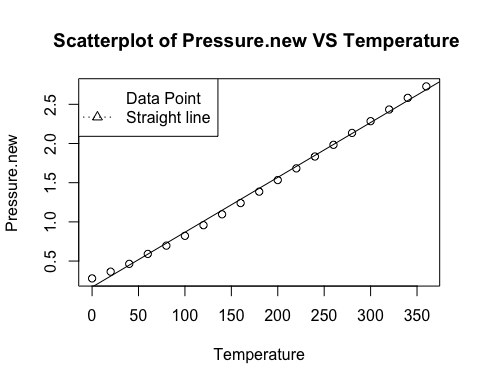
#(b)

pressure$pressure.new<-pressure$pressure^(3/20)  
plot(x=pressure$temperature,y=pressure$pressure.new,main="Scatterplot of Pressure.new VS Temperature ",xlab="Temperature",ylab="Pressure.new")  
abline(0.168,0.007)



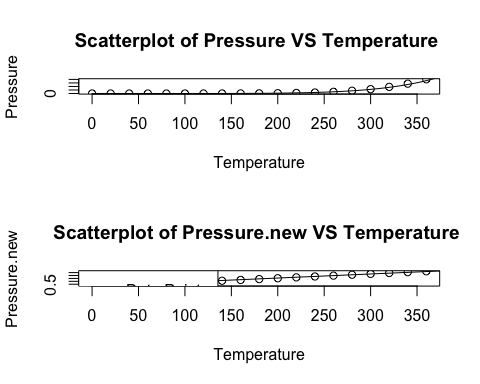
#(c)

pressure$pressure.new<-pressure$pressure^(3/20)  
plot(x=pressure$temperature,y=pressure$pressure.new,main="Scatterplot of Pressure.new VS Temperature ",xlab="Temperature",ylab="Pressure.new")  
abline(0.168,0.007)  
legend("topleft",legend = c("Data Point","Straight line"),lty = c(NA,3),pch = c(NA,2))



#(d)

#2x1 layout  
par(mfrow=c(2,1))  
#1st plot  
plot(x=pressure$temperature,y=pressure$pressure,main="Scatterplot of Pressure VS Temperature ",xlab="Temperature",ylab="Pressure")  
curve((0.168+0.007\*x)^(20/3),from = 0,to=400,add=TRUE)  
#2nd plot  
pressure$pressure.new<-pressure$pressure^(3/20)  
plot(x=pressure$temperature,y=pressure$pressure.new,main="Scatterplot of Pressure.new VS Temperature ",xlab="Temperature",ylab="Pressure.new")  
abline(0.168,0.007)  
legend("topleft",legend = c("Data Point","Straight line"),lty = c(NA,3),pch = c(NA,2))



#1x2 layout  
par(mfrow=c(1,2))  
#1st plot  
plot(x=pressure$temperature,y=pressure$pressure,main="Scatterplot of Pressure VS Temperature ",xlab="Temperature",ylab="Pressure")  
curve((0.168+0.007\*x)^(20/3),from = 0,to=400,add=TRUE)  
#2nd plot  
pressure$pressure.new<-pressure$pressure^(3/20)  
plot(x=pressure$temperature,y=pressure$pressure.new,main="Scatterplot of Pressure.new VS Temperature ",xlab="Temperature",ylab="Pressure.new")  
abline(0.168,0.007)  
legend("topleft",legend = c("Data Point","Straight line"),lty = c(NA,3),pch = c(NA,2))

