LAB 3\_2

setwd(".")  
library(knitr)  
library(ggplot2)  
library(sqldf)

## Loading required package: gsubfn

## Loading required package: proto

## Loading required package: RSQLite

library(DMwR2)

## Registered S3 method overwritten by 'xts':  
## method from  
## as.zoo.xts zoo

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

default<-read.table("Archivos/Puromicina.txt", dec = ".", sep=",",header=T)

#APARTADO A #Media, mediana

attach(default)  
media\_VR<-aggregate(velocidad\_reaccion~Puromicina,default,mean)  
kable(media\_VR)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 141.5833 |
| untreated | 110.7273 |

mediana\_VR<-aggregate(velocidad\_reaccion~Puromicina,default,median)  
kable(mediana\_VR)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 145.5 |
| untreated | 115.0 |
| #APARTADO B |  |
| #parámetros d | e dispersión |

varianza\_VR<-aggregate(velocidad\_reaccion~Puromicina,default,var)  
kable(varianza\_VR)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 2805.356 |
| untreated | 1334.218 |

sd\_VR<-aggregate(velocidad\_reaccion~Puromicina,default,sd)  
kable(sd\_VR)

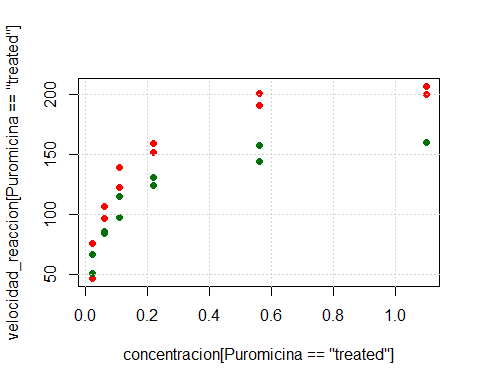
|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 52.96561 |
| untreated | 36.52695 |

detach()

#APARTADO C y D #Visualizar si la concetraciónde sustrato influye

#Datos sin ajuste.

attach(default)  
plot(concentracion[Puromicina=="treated"], velocidad\_reaccion[Puromicina=="treated"],   
 col="red", pch=19)  
points(concentracion[Puromicina=="untreated"], velocidad\_reaccion[Puromicina=="untreated"],   
 col="#037005", pch=19)  
grid()



detach()

#Datos ajustados.

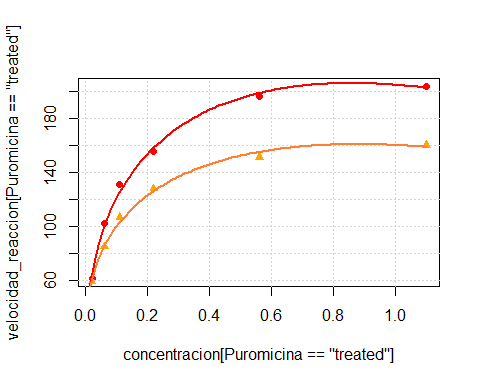
medias\_datos<-aggregate(velocidad\_reaccion~Puromicina+concentracion, default, mean)  
kable(medias\_datos)

|  |  |  |
| --- | --- | --- |
| Puromicina | concentracion | velocidad\_reaccion |
| treated | 0.02 | 61.5 |
| untreated | 0.02 | 59.0 |
| treated | 0.06 | 102.0 |
| untreated | 0.06 | 85.0 |
| treated | 0.11 | 131.0 |
| untreated | 0.11 | 106.5 |
| treated | 0.22 | 155.5 |
| untreated | 0.22 | 127.5 |
| treated | 0.56 | 196.0 |
| untreated | 0.56 | 151.0 |
| treated | 1.10 | 203.5 |
| untreated | 1.10 | 160.0 |

attach(medias\_datos)  
plot(concentracion[Puromicina=="treated"],   
 velocidad\_reaccion[Puromicina=="treated"],   
 col="red", pch=19)  
grid()  
points(concentracion[Puromicina=="untreated"],   
 velocidad\_reaccion[Puromicina=="untreated"],   
 col="orange", pch=17)  
  
datos\_ord\_velocidad<-default[order(default$velocidad\_reaccion),]  
kable(datos\_ord\_velocidad)

|  |  |  |  |
| --- | --- | --- | --- |
|  | concentracion | velocidad\_reaccion | Puromicina |
| 2 | 0.02 | 47 | treated |
| 14 | 0.02 | 51 | untreated |
| 13 | 0.02 | 67 | untreated |
| 1 | 0.02 | 76 | treated |
| 15 | 0.06 | 84 | untreated |
| 16 | 0.06 | 86 | untreated |
| 3 | 0.06 | 97 | treated |
| 17 | 0.11 | 98 | untreated |
| 4 | 0.06 | 107 | treated |
| 18 | 0.11 | 115 | untreated |
| 5 | 0.11 | 123 | treated |
| 20 | 0.22 | 124 | untreated |
| 19 | 0.22 | 131 | untreated |
| 6 | 0.11 | 139 | treated |
| 21 | 0.56 | 144 | untreated |
| 8 | 0.22 | 152 | treated |
| 22 | 0.56 | 158 | untreated |
| 7 | 0.22 | 159 | treated |
| 23 | 1.10 | 160 | untreated |
| 9 | 0.56 | 191 | treated |
| 12 | 1.10 | 200 | treated |
| 10 | 0.56 | 201 | treated |
| 11 | 1.10 | 207 | treated |

yt<-velocidad\_reaccion[Puromicina=="treated"]  
xt<-concentracion[Puromicina=="treated"]  
modelo1<-lm(yt~xt+I(xt^(1/2)))  
xv<-seq(from=0,to=1.1,by=0.01)  
yv<-predict(modelo1,list(xt=xv))  
lines(xv,yv, col="red", lwd=2)  
  
#No funciona  
yt2<-velocidad\_reaccion[Puromicina=="untreated"]  
xt2<-concentracion[Puromicina=="untreated"]  
modelo2<-lm(yt2~xt2+I(xt2^(1/2)))  
xv2<-seq(from=0, to=1.1, by=0.01)  
yv2<-predict(modelo2, list(xt2=xv2))  
lines(xv2,yv2, col="#ff8033", lwd=2)



detach()

# APARTADO F

datos\_NA<-read.table("Archivos/Puromicina\_NA.txt", dec = ".", sep=",", header=T)  
kable(datos\_NA)

|  |  |  |
| --- | --- | --- |
| concentracion | velocidad\_reaccion | Puromicina |
| 0.02 | 76 | treated |
| 0.02 | 47 | treated |
| 0.06 | 97 | treated |
| 0.06 | 107 | treated |
| NA | 123 | treated |
| 0.11 | 139 | treated |
| 0.22 | 159 | treated |
| 0.22 | 152 | treated |
| 0.56 | 191 | NA |
| 0.56 | 201 | treated |
| 1.10 | NA | treated |
| 1.10 | 200 | treated |
| 0.02 | 67 | untreated |
| 0.02 | 51 | untreated |
| 0.06 | 84 | untreated |
| 0.06 | 86 | untreated |
| 0.11 | 98 | untreated |
| 0.11 | 115 | NA |
| 0.22 | 131 | untreated |
| 0.22 | NA | untreated |
| 0.56 | 144 | untreated |
| 0.56 | NA | untreated |
| 1.10 | 160 | untreated |

# Completa los NA con un valor centralizado.  
datos\_ajustados<-centralImputation(datos\_NA)  
kable(datos\_ajustados)

|  |  |  |
| --- | --- | --- |
| concentracion | velocidad\_reaccion | Puromicina |
| 0.020 | 76 | treated |
| 0.020 | 47 | treated |
| 0.060 | 97 | treated |
| 0.060 | 107 | treated |
| 0.165 | 123 | treated |
| 0.110 | 139 | treated |
| 0.220 | 159 | treated |
| 0.220 | 152 | treated |
| 0.560 | 191 | treated |
| 0.560 | 201 | treated |
| 1.100 | 119 | treated |
| 1.100 | 200 | treated |
| 0.020 | 67 | untreated |
| 0.020 | 51 | untreated |
| 0.060 | 84 | untreated |
| 0.060 | 86 | untreated |
| 0.110 | 98 | untreated |
| 0.110 | 115 | treated |
| 0.220 | 131 | untreated |
| 0.220 | 119 | untreated |
| 0.560 | 144 | untreated |
| 0.560 | 119 | untreated |
| 1.100 | 160 | untreated |

# Método del vecino mas cercano. Ajusta los NA al k mas cercano a ellos.  
datos\_ajustados\_vecino<-knnImputation(datos\_NA)  
kable(datos\_ajustados\_vecino)

|  |  |  |
| --- | --- | --- |
| concentracion | velocidad\_reaccion | Puromicina |
| 0.0200000 | 76.0000 | treated |
| 0.0200000 | 47.0000 | treated |
| 0.0600000 | 97.0000 | treated |
| 0.0600000 | 107.0000 | treated |
| 0.1506064 | 123.0000 | treated |
| 0.1100000 | 139.0000 | treated |
| 0.2200000 | 159.0000 | treated |
| 0.2200000 | 152.0000 | treated |
| 0.5600000 | 191.0000 | treated |
| 0.5600000 | 201.0000 | treated |
| 1.1000000 | 175.9734 | treated |
| 1.1000000 | 200.0000 | treated |
| 0.0200000 | 67.0000 | untreated |
| 0.0200000 | 51.0000 | untreated |
| 0.0600000 | 84.0000 | untreated |
| 0.0600000 | 86.0000 | untreated |
| 0.1100000 | 98.0000 | untreated |
| 0.1100000 | 115.0000 | treated |
| 0.2200000 | 131.0000 | untreated |
| 0.2200000 | 106.0601 | untreated |
| 0.5600000 | 144.0000 | untreated |
| 0.5600000 | 126.6490 | untreated |
| 1.1000000 | 160.0000 | untreated |

# Empleando NA.omit()  
datos\_ajustados\_naomit<-na.omit(datos\_NA)  
kable(datos\_ajustados\_naomit)

|  |  |  |  |
| --- | --- | --- | --- |
|  | concentracion | velocidad\_reaccion | Puromicina |
| 1 | 0.02 | 76 | treated |
| 2 | 0.02 | 47 | treated |
| 3 | 0.06 | 97 | treated |
| 4 | 0.06 | 107 | treated |
| 6 | 0.11 | 139 | treated |
| 7 | 0.22 | 159 | treated |
| 8 | 0.22 | 152 | treated |
| 10 | 0.56 | 201 | treated |
| 12 | 1.10 | 200 | treated |
| 13 | 0.02 | 67 | untreated |
| 14 | 0.02 | 51 | untreated |
| 15 | 0.06 | 84 | untreated |
| 16 | 0.06 | 86 | untreated |
| 17 | 0.11 | 98 | untreated |
| 19 | 0.22 | 131 | untreated |
| 21 | 0.56 | 144 | untreated |
| 23 | 1.10 | 160 | untreated |

# Empleado complete.cases()  
datos\_ajustados\_cc<-datos\_NA[complete.cases(datos\_NA)==TRUE,]  
kable(datos\_ajustados\_cc)

|  |  |  |  |
| --- | --- | --- | --- |
|  | concentracion | velocidad\_reaccion | Puromicina |
| 1 | 0.02 | 76 | treated |
| 2 | 0.02 | 47 | treated |
| 3 | 0.06 | 97 | treated |
| 4 | 0.06 | 107 | treated |
| 6 | 0.11 | 139 | treated |
| 7 | 0.22 | 159 | treated |
| 8 | 0.22 | 152 | treated |
| 10 | 0.56 | 201 | treated |
| 12 | 1.10 | 200 | treated |
| 13 | 0.02 | 67 | untreated |
| 14 | 0.02 | 51 | untreated |
| 15 | 0.06 | 84 | untreated |
| 16 | 0.06 | 86 | untreated |
| 17 | 0.11 | 98 | untreated |
| 19 | 0.22 | 131 | untreated |
| 21 | 0.56 | 144 | untreated |
| 23 | 1.10 | 160 | untreated |

attach(datos\_ajustados\_naomit)  
media\_VR\_naomit<-aggregate(velocidad\_reaccion~Puromicina,datos\_ajustados\_naomit,mean)  
kable(media\_VR\_naomit)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 130.8889 |
| untreated | 102.6250 |

mediana\_VR\_naomit<-aggregate(velocidad\_reaccion~Puromicina,datos\_ajustados\_naomit,median)  
kable(mediana\_VR\_naomit)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 139 |
| untreated | 92 |

detach()  
  
attach(datos\_ajustados\_cc)  
media\_VR\_cc<-aggregate(velocidad\_reaccion~Puromicina,datos\_ajustados\_cc,mean)  
kable(media\_VR\_cc)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 130.8889 |
| untreated | 102.6250 |

mediana\_VR\_cc<-aggregate(velocidad\_reaccion~Puromicina,datos\_ajustados\_cc,median)  
kable(mediana\_VR\_cc)

|  |  |
| --- | --- |
| Puromicina | velocidad\_reaccion |
| treated | 139 |
| untreated | 92 |

detach()  
  
#Creramos una tabl nueva con 0  
datos\_ceros<-datos\_NA  
datos\_ceros$concentracion[is.na(datos\_ceros$concentracion)]<-0  
datos\_ceros$velocidad\_reaccion[is.na(datos\_ceros$velocidad\_reaccion)]<-0  
kable(datos\_ceros)

|  |  |  |
| --- | --- | --- |
| concentracion | velocidad\_reaccion | Puromicina |
| 0.02 | 76 | treated |
| 0.02 | 47 | treated |
| 0.06 | 97 | treated |
| 0.06 | 107 | treated |
| 0.00 | 123 | treated |
| 0.11 | 139 | treated |
| 0.22 | 159 | treated |
| 0.22 | 152 | treated |
| 0.56 | 191 | NA |
| 0.56 | 201 | treated |
| 1.10 | 0 | treated |
| 1.10 | 200 | treated |
| 0.02 | 67 | untreated |
| 0.02 | 51 | untreated |
| 0.06 | 84 | untreated |
| 0.06 | 86 | untreated |
| 0.11 | 98 | untreated |
| 0.11 | 115 | NA |
| 0.22 | 131 | untreated |
| 0.22 | 0 | untreated |
| 0.56 | 144 | untreated |
| 0.56 | 0 | untreated |
| 1.10 | 160 | untreated |

datos\_ceros\_knn<-knnImputation(datos\_ceros)  
  
# Resumenes de las tablas ajustadas  
summary(datos\_NA)

## concentracion velocidad\_reaccion Puromicina  
## Min. :0.0200 Min. : 47.0 treated :11   
## 1st Qu.:0.0600 1st Qu.: 85.5 untreated:10   
## Median :0.1650 Median :119.0 NA's : 2   
## Mean :0.3214 Mean :121.4   
## 3rd Qu.:0.5600 3rd Qu.:153.8   
## Max. :1.1000 Max. :201.0   
## NA's :1 NA's :3

summary(datos\_ajustados)

## concentracion velocidad\_reaccion Puromicina  
## Min. :0.0200 Min. : 47.0 treated :13   
## 1st Qu.:0.0600 1st Qu.: 91.5 untreated:10   
## Median :0.1650 Median :119.0   
## Mean :0.3146 Mean :121.1   
## 3rd Qu.:0.5600 3rd Qu.:148.0   
## Max. :1.1000 Max. :201.0

summary(datos\_ajustados\_vecino)

## concentracion velocidad\_reaccion Puromicina  
## Min. :0.0200 Min. : 47.0 treated :13   
## 1st Qu.:0.0600 1st Qu.: 91.5 untreated:10   
## Median :0.1506 Median :123.0   
## Mean :0.3139 Mean :123.3   
## 3rd Qu.:0.5600 3rd Qu.:155.5   
## Max. :1.1000 Max. :201.0

summary(datos\_ajustados\_naomit)

## concentracion velocidad\_reaccion Puromicina  
## Min. :0.0200 Min. : 47.0 treated :9   
## 1st Qu.:0.0600 1st Qu.: 84.0 untreated:8   
## Median :0.1100 Median :107.0   
## Mean :0.2659 Mean :117.6   
## 3rd Qu.:0.2200 3rd Qu.:152.0   
## Max. :1.1000 Max. :201.0

summary(datos\_ajustados\_cc)

## concentracion velocidad\_reaccion Puromicina  
## Min. :0.0200 Min. : 47.0 treated :9   
## 1st Qu.:0.0600 1st Qu.: 84.0 untreated:8   
## Median :0.1100 Median :107.0   
## Mean :0.2659 Mean :117.6   
## 3rd Qu.:0.2200 3rd Qu.:152.0   
## Max. :1.1000 Max. :201.0

summary(datos\_ceros\_knn)

## concentracion velocidad\_reaccion Puromicina  
## Min. :0.0000 Min. : 0.0 treated :13   
## 1st Qu.:0.0600 1st Qu.: 71.5 untreated:10   
## Median :0.1100 Median :107.0   
## Mean :0.3074 Mean :105.6   
## 3rd Qu.:0.5600 3rd Qu.:148.0   
## Max. :1.1000 Max. :201.0