Consumo de alcohol en los estudiantes

Gerson Omar Pérez & Victor Cortez Sis

27/5/2021

## Enlace del set de datos

<https://www.kaggle.com/uciml/student-alcohol-consumption?select=student-por.csv>

## Instalacion de paquetes y librerias

install.packages(“dplyr”) install.packages(“ggplot2”) install.packages(“corrplot”) require(stats) library(dplyr) library(ggplot) library(ggplot2) library(corrplot)

## Cargar el data set

alcohol <- read.csv("student-por.csv")

## 1

## ¿La relacion que tiene con su familia tiene relacion con en tiemmpo que el estudiante pasa fuera de casa?

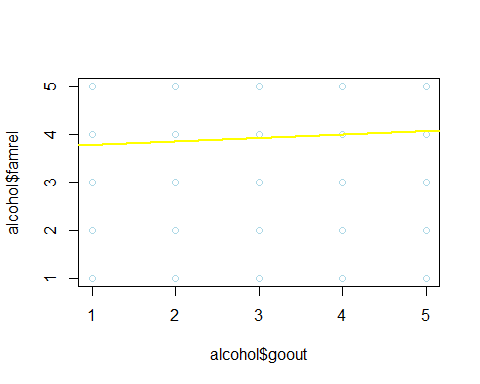
## Ho: El tiempo que pasa fuera de casa SI es causado por la relacion familiar del estudiante

## H1: El tiempo que pasa fuera de casa NO es causado por la relacion familiar del estudiante

cor(alcohol$famrel, alcohol$goout)

## [1] 0.08970657

plot(alcohol$goout, alcohol$famrel, col="lightblue")  
abline(lm(alcohol$famrel ~ alcohol$goout), col = "yellow", lwd = 2)



#Correlacion: 0.08970657

## Los datos poseen una correlacion negativa, por lo tanto deducimos que el tiempo que el estudiante pasa fuera de casa no esta conectado con la relacion que el estudiante tiene en su nucleo familiar.

## 2

## ¿El genero del estudiante es un determinante para que este tenga una enseñanza de alta calidad?

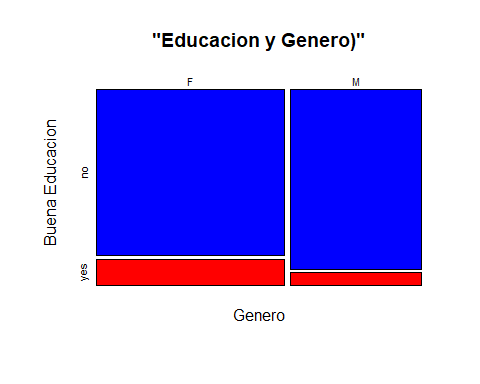
## Ho: Es probable que el genero NO garantice una enseñanza de alta calidad

## H1: Es probable que el genero SI garantice una enseñanza de alta calidad

tablaa<-table(alcohol$sex, alcohol$schoolsup, dnn = c("Genero","Buena Educacion"))  
tablaa

## Buena Educacion  
## Genero no yes  
## F 332 51  
## M 249 17

mosaicplot(tablaa, main = deparse("Educacion y Genero)"), col=c("blue","red","yellow"))



chisq.test(tablaa)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: tablaa  
## X-squared = 7.3042, df = 1, p-value = 0.006879

0.006879

## [1] 0.006879

## Se rechaza la hipotesis nula debido a que nuestro valor de chi cuadrado es 0.006879, con lo cual podemos decir que es probable que se pueda obtener una educacion de alta calidad segun el genero del joven estudiante

## 3

## ¿La relación que tiene dentro de su hogar influye en los fracasos que ha tenido en su formación académica?

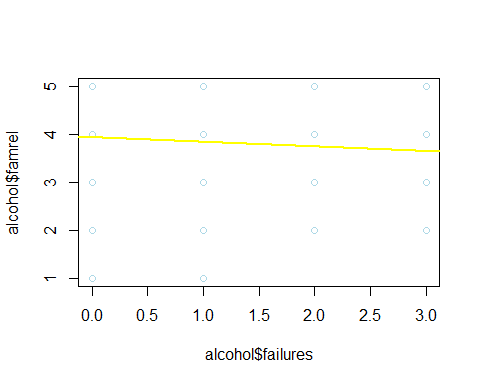
## Ho: La relacion familiar si es la posible causante de sus fracasos estudiantiles

## H1: La relacion familiar no es la posible causante de sus fracasos estudiantiles

cor(alcohol$famrel, alcohol$failures)

## [1] -0.06264516

plot(alcohol$failures, alcohol$famrel, col="lightblue")  
abline(lm(alcohol$famrel ~ alcohol$failures), col = "yellow", lwd = 2)



#Correlacion: -0.06264516

## Los datos obtenidos nos dan una correlacion negativa, debido a esto podemos decir que la relacion familiar no es la causante de sus fracasos como estudiante

## 4

## ¿EL hecho de tener una familia de alta calidad es algo que le otorga estudios de alta calidad al estudiante?

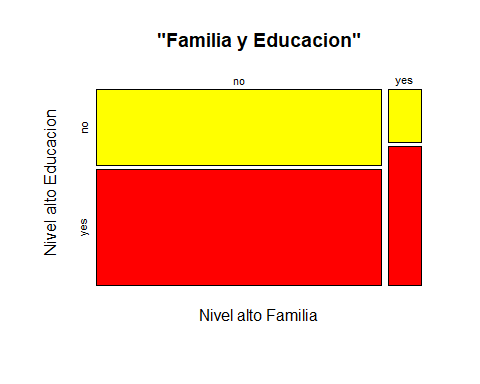
## Ho: Tener una buena familia NO es un factor para tener educacion de alta calidad

## H1: Tener una buena familia SI es un factor para tener educacion de alta calidad

tablae <- table(alcohol$schoolsup, alcohol$famsup, dnn=c("Nivel alto Familia","Nivel alto Educacion" ))  
tablae

## Nivel alto Educacion  
## Nivel alto Familia no yes  
## no 232 349  
## yes 19 49

mosaicplot(tablae, main=deparse("Familia y Educacion"), col=c("yellow","red","blue"))



chisq.test(tablae)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: tablae  
## X-squared = 3.2016, df = 1, p-value = 0.07356

0.07356<0.05

## [1] FALSE

## No se rechaza la hipotesis nula debido a que nuestro valor de chi cuadrado es de 0.07356, por lo tanto podemos definir que el hecho de tener una familia de alta calidad es algo que puede llegar a garantizar que un estudiante tenga acceso a una educacion de alto nivel

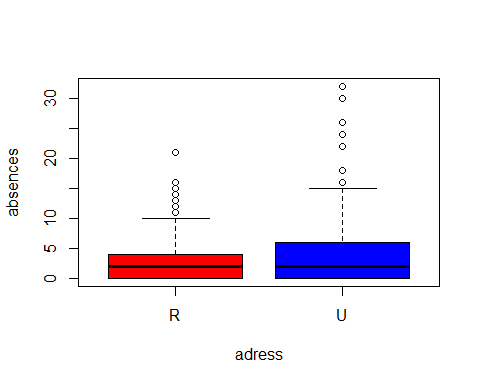
## 5

## La distancia del casa al colegio (direccion) tiene algo que ver con las ausencias que ha tenido el estudiante en su respectiva escuela?

## Ho: La distancia de casa-colegio SI influye en las ausencias que ha tenido

## H1: La distancia de casa-colgeio NO influye en las ausencias que ha tenido

tablacontinengcia<- table(alcohol$adress, alcohol$absences)  
boxplot(absences ~ adress, data=alcohol, col=c("red", "blue"))



chisq.test(tablacontinengcia)

## Warning in chisq.test(tablacontinengcia): Chi-squared approximation may be  
## incorrect

##   
## Pearson's Chi-squared test  
##   
## data: tablacontinengcia  
## X-squared = 30.913, df = 23, p-value = 0.1249

0.1249<0.05

## [1] FALSE

## No se rechaza la hipotesis nula debido a que nuestro valor de chi cuadrado es 0.1249, por lo tanto se determina de manera oficial que la distancia que hay entre la casa y el colegio, si es un factor que provoca que los estudiantes sean ausentes en sus clases

## 6

## El estado personal del estudiante se ve afectado por la persona que esta a cargo de la custodia del joven?

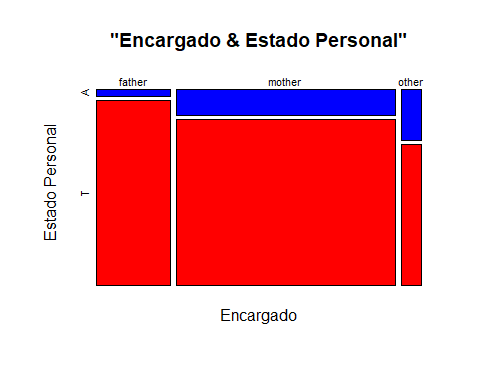
## Ho: La persona encargada NO afecta el estado personal

## H1: La persona encargada SI afecta el estado personal

tablai <- table(alcohol$guardian, alcohol$Pstatus, dnn=c("Encargado","Estado Personal" ))  
tablai

## Estado Personal  
## Encargado A T  
## father 6 147  
## mother 63 392  
## other 11 30

mosaicplot(tablai, main=deparse("Encargado & Estado Personal"), col=c("blue","red","yellow"))



chisq.test(tablai)

##   
## Pearson's Chi-squared test  
##   
## data: tablai  
## X-squared = 18.953, df = 2, p-value = 7.664e-05

7.664e-05<0.05

## [1] TRUE

## Se rechaza la hipotesis nula debido a que nuestro chi cuadrado es 7.664e, por esta razon se dice que la persona que este a cargo del joven estudiante si puede lelgar a tener influecia en el estado personal del joven

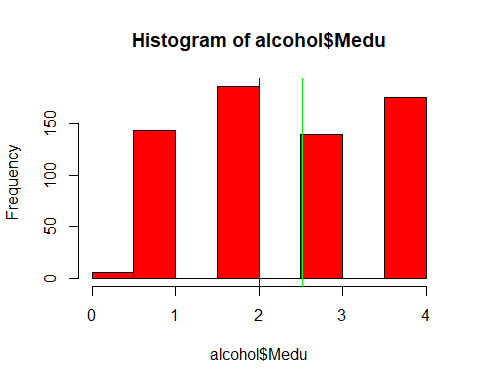
## 7

## ¿Cual es la media del nivel de educacion que tiene un hombre?

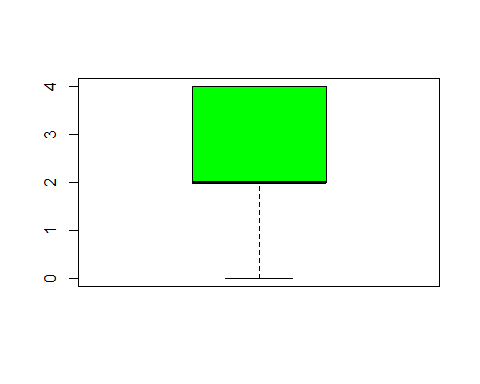
## Ho: La media del nivel de educacion para un hombre es mu=4

## H1: La media del nivel de educacion para un hombre es mu=/=4

hist(alcohol$Medu, col="red")  
mediaMedu <- mean(alcohol$Medu)  
abline(v=mediaMedu, col="green")  
medianaMedu <- median(alcohol$Medu)  
abline (v=medianaMedu, col="blue")



boxplot(alcohol$Medu, col="green")



t.test(alcohol$Medu, mu=4, alternative="two.sided")

##   
## One Sample t-test  
##   
## data: alcohol$Medu  
## t = -33.353, df = 648, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 4  
## 95 percent confidence interval:  
## 2.427187 2.602088  
## sample estimates:  
## mean of x   
## 2.514638

#Media: 2.51  
  
2.2e-16<0.05

## [1] TRUE

## Se rechaza la hipotesis nula, lo cual nos indica que la media del nivel de educacion de un hombre es diferente a 4

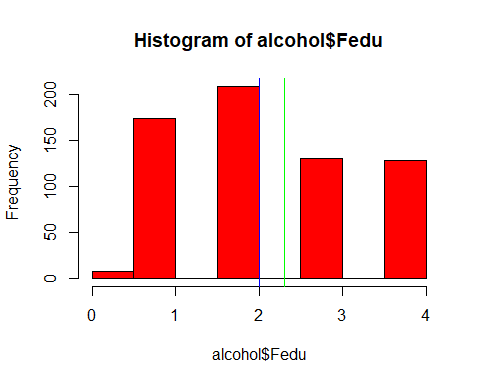
## 8

## ¿Cual es la media del nivel de educacion que tiene una mujer?

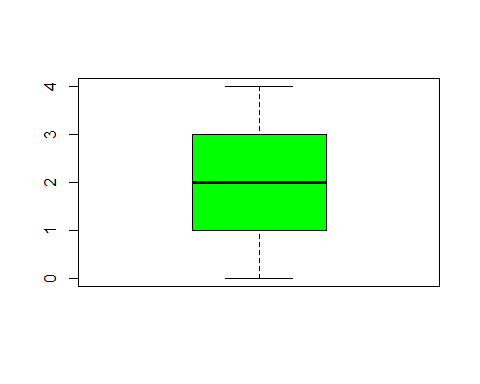
## Ho: La media del nivel de educacion para una mujer es mu=4

## H1: La media del nivel de educacion para una mujer es mu=/=4

hist(alcohol$Fedu, col="red")  
mediaFedu <- mean(alcohol$Fedu)  
abline(v=mediaFedu, col="green")  
medianaFedu <- median(alcohol$Fedu)  
abline (v=medianaFedu, col="blue")



boxplot(alcohol$Fedu, col="green")



t.test(alcohol$Fedu, mu=4, alternative="two.sided")

##   
## One Sample t-test  
##   
## data: alcohol$Fedu  
## t = -39.22, df = 648, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 4  
## 95 percent confidence interval:  
## 2.221844 2.391407  
## sample estimates:  
## mean of x   
## 2.306626

#Media: 2.30  
  
2.2e-16<0.05

## [1] TRUE

## Se rechaza la hipotesis nula, lo cual nos indica que la media del nivel de educacion de una mujer es diferente a 4

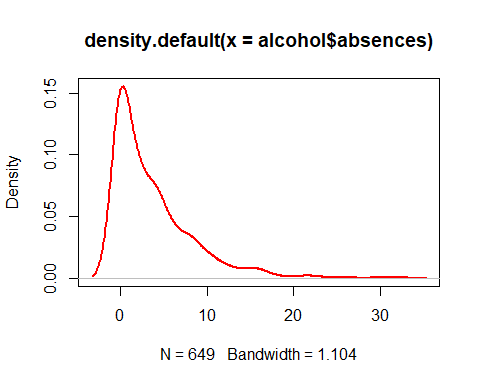
## 9

## Nivel de usencia en el colegio de los estudiantes

## Ho: Si hay normalidad en los datos

## H1: NO hay normalidad en los datos

plot(density(alcohol$absences), col="red", lwd=2)



shapiro.test(alcohol$absences)

##   
## Shapiro-Wilk normality test  
##   
## data: alcohol$absences  
## W = 0.77174, p-value < 2.2e-16

2.2e-16<0.05

## [1] TRUE

## Se rechaza la hipotesis nula, por lo tanto esto nos indica que los datos analizados en esta ocasion no tienen normalidad

## Tambien se nos muestra que los datos estan sesgados a la derecha

## 10

## ¿Existe diferencia en la calidad de enseñanza que puede tener un hombre y una mujer?

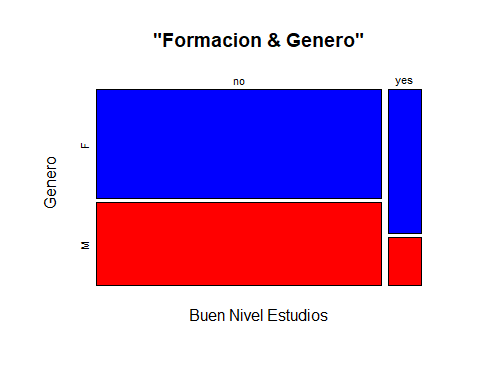
## Ho: No hay tal diferencia en la calidad de enseñanza con relacion al genero

## H1: Si hay tal diferencia en la calidad de enseñanza con relacion al genero

tablao <- table(alcohol$schoolsup, alcohol$sex, dnn=c("Buen Nivel Estudios","Genero" ))  
tablao

## Genero  
## Buen Nivel Estudios F M  
## no 332 249  
## yes 51 17

mosaicplot(tablao, main=deparse("Formacion & Genero"), col=c("blue","red","yellow"))



chisq.test(tablao)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: tablao  
## X-squared = 7.3042, df = 1, p-value = 0.006879

0.006879<0.05

## [1] TRUE

## Se rechaza la hipotesis nula ya que el pvalue es menor al nivel de significancia, esto nos da a entender que el el genero es una variable determinante que le garantiza al estudiante una buena formacion academica

## 11

## ¿Que correlacion tiene el nivel de estudios de un hombre con respecto a los fracasos que ha tenido en su vida estudiantil?

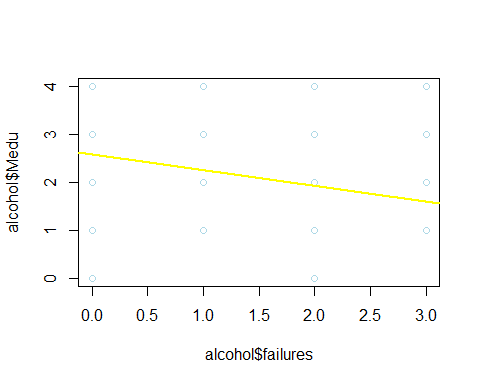
## Ho: El nivel de estudios si esta correlacionado con los fracasos

## H1: El nivel de estudios no esta correlacionado con los fracasos

cor(alcohol$Medu, alcohol$failures)

## [1] -0.1722103

plot(alcohol$failures, alcohol$Medu, col="lightblue")  
abline(lm(alcohol$Medu ~ alcohol$failures), col = "yellow", lwd = 2)



#Correlacion entre las dos variables: -0.1722103

## La correlacion es negativa, por lo tanto se llega a la conclusion de que el nivel de estudios que tienen los hombres no esta relacionado respecto a los fracasos que tiene en el colegio o escuela.

## 12

## ¿Que correlacion tiene el nivel de estudios de una mujer con respecto a los fracasos que ha tenido?

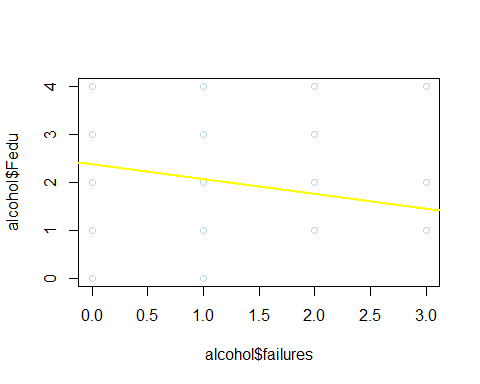
## Ho: El nivel de estudios si esta correlacionado con los fracasos

## H1: El nivel de estudios no esta correlacionado con los fracasos

cor(alcohol$Fedu, alcohol$failures)

## [1] -0.165915

plot(alcohol$failures, alcohol$Fedu, col="lightblue")  
abline(lm(alcohol$Fedu ~ alcohol$failures), col = "yellow", lwd = 2)



#Correlacion entre las dos variables: -0.165915

## La correlacion es negativa, por lo tanto se llega a la conclusion de que el nivel de estudios que tienen las mujeres no esta relacionado respecto a los fracasos que han tenido

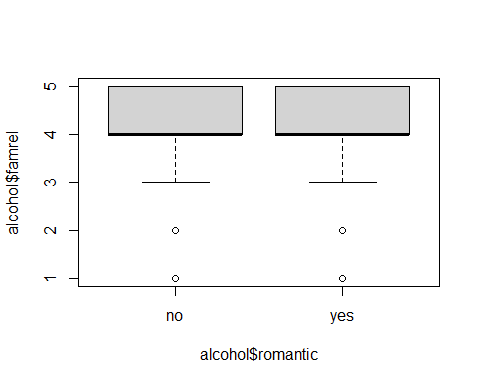
## 13

## La relacion que el estudiente tiene dentro de su familia afecta el lado amoroso/empatico que posee en su vida estudiantil?

## Ho: No existe diferencia en la relacion familiar con respecto al nivel de cariño que posee

## H1: Si existe diferencia en la relacion familiar con respecto al nivel de cariño que posee

boxplot(alcohol$famrel~alcohol$romantic)



t.test(alcohol$famrel~alcohol$romantic)

##   
## Welch Two Sample t-test  
##   
## data: alcohol$famrel by alcohol$romantic  
## t = 1.1152, df = 460.15, p-value = 0.2653  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.06777751 0.24565281  
## sample estimates:  
## mean in group no mean in group yes   
## 3.963415 3.874477

0.2653<0.05

## [1] FALSE

## Se rechaza la hipotesis nula debido a que el valor de t.test es 0.2653, entonces con esto se puede decir que la relacion familiar es algo que si afecta el nivel de cariño/amor que el estudiante muestre dentro de su entorno estudiantil

## 14

## ¿Una educacion de alta calidad garantiza que esl estudiante pueda sacar el diversificado(preparatoria)?

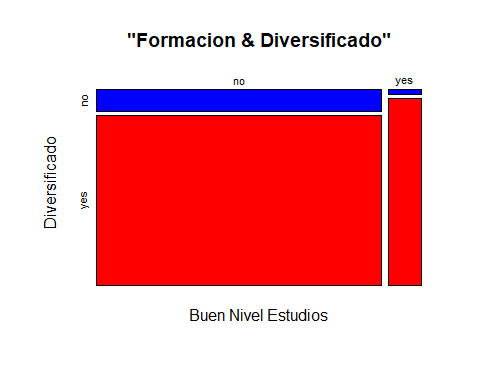
## Ho: Una buena formacion academica no garantiza que el estudiante pueda estudiar el diversificado

## H1: Una buena formacion academica si garantiza que el estudiante pueda estudiar el diversificado

tablau <- table(alcohol$schoolsup, alcohol$higher, dnn=c("Buen Nivel Estudios","Diversificado" ))  
tablau

## Diversificado  
## Buen Nivel Estudios no yes  
## no 67 514  
## yes 2 66

mosaicplot(tablau, main=deparse("Formacion & Diversificado"), col=c("blue","red","yellow"))



chisq.test(tablau)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: tablau  
## X-squared = 3.8674, df = 1, p-value = 0.04923

0.04923<0.05

## [1] TRUE

## Se rechaza la hipotesis nula ya que el chi cuadrado es 0.04923, por lo tanto determinamos que un alto nivel de estudios si puede llegar a garantizar que el estudiante curse el diversificado

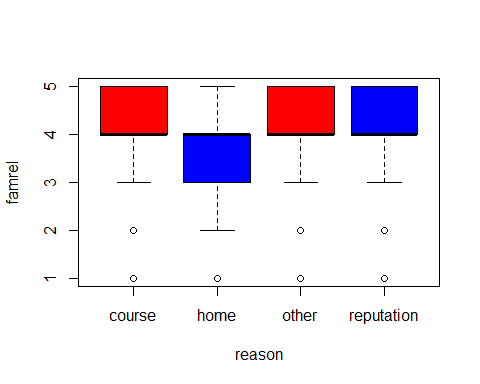
## 15

## ¿El motivo del alcoholismo del joven estudiante puede estar siendo causado por la relacion familiar que se vive en su hogar?

## Ho: La relacion familiar podria ser la causante del alcoholismo del estudiante

## H1: La relacion familiar podria no ser la causante del alcoholismo dle estudiante

tablacontinengcia<- table(alcohol$reason, alcohol$famrel)  
boxplot(famrel ~ reason, data=alcohol, col=c("red", "blue"))



chisq.test(tablacontinengcia)

## Warning in chisq.test(tablacontinengcia): Chi-squared approximation may be  
## incorrect

##   
## Pearson's Chi-squared test  
##   
## data: tablacontinengcia  
## X-squared = 9.8196, df = 12, p-value = 0.6318

0.6318<0.05

## [1] FALSE

## No se rechaza la hipotesis nula devico a que el chi cuadrado es 0.6318, por lo tanto se puede decir que el motivo de alcoholismo del estudiante si puede estar siendo ocasionado por la relaion que tiene dentro de su hogar

## Resumen estadistico de cada una de las variables

summary(alcohol)

## school sex age adress   
## Length:649 Length:649 Min. :15.00 Length:649   
## Class :character Class :character 1st Qu.:16.00 Class :character   
## Mode :character Mode :character Median :17.00 Mode :character   
## Mean :16.74   
## 3rd Qu.:18.00   
## Max. :22.00   
## famsize Pstatus Medu Fedu   
## Length:649 Length:649 Min. :0.000 Min. :0.000   
## Class :character Class :character 1st Qu.:2.000 1st Qu.:1.000   
## Mode :character Mode :character Median :2.000 Median :2.000   
## Mean :2.515 Mean :2.307   
## 3rd Qu.:4.000 3rd Qu.:3.000   
## Max. :4.000 Max. :4.000   
## Mjob Fjob reason guardian   
## Length:649 Length:649 Length:649 Length:649   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## traveltime studytime failures schoolsup   
## Min. :1.000 Min. :1.000 Min. :0.0000 Length:649   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.0000 Class :character   
## Median :1.000 Median :2.000 Median :0.0000 Mode :character   
## Mean :1.569 Mean :1.931 Mean :0.2219   
## 3rd Qu.:2.000 3rd Qu.:2.000 3rd Qu.:0.0000   
## Max. :4.000 Max. :4.000 Max. :3.0000   
## famsup paid activities nursery   
## Length:649 Length:649 Length:649 Length:649   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## higher internet romantic famrel   
## Length:649 Length:649 Length:649 Min. :1.000   
## Class :character Class :character Class :character 1st Qu.:4.000   
## Mode :character Mode :character Mode :character Median :4.000   
## Mean :3.931   
## 3rd Qu.:5.000   
## Max. :5.000   
## freetime goout Dalc Walc health   
## Min. :1.00 Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000   
## 1st Qu.:3.00 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:2.000   
## Median :3.00 Median :3.000 Median :1.000 Median :2.00 Median :4.000   
## Mean :3.18 Mean :3.185 Mean :1.502 Mean :2.28 Mean :3.536   
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:2.000 3rd Qu.:3.00 3rd Qu.:5.000   
## Max. :5.00 Max. :5.000 Max. :5.000 Max. :5.00 Max. :5.000   
## absences G1 G2 G3   
## Min. : 0.000 Min. : 0.0 Min. : 0.00 Min. : 0.00   
## 1st Qu.: 0.000 1st Qu.:10.0 1st Qu.:10.00 1st Qu.:10.00   
## Median : 2.000 Median :11.0 Median :11.00 Median :12.00   
## Mean : 3.659 Mean :11.4 Mean :11.57 Mean :11.91   
## 3rd Qu.: 6.000 3rd Qu.:13.0 3rd Qu.:13.00 3rd Qu.:14.00   
## Max. :32.000 Max. :19.0 Max. :19.00 Max. :19.00

## Estructura de las variables

str(alcohol)

## 'data.frame': 649 obs. of 33 variables:  
## $ school : chr "GP" "GP" "GP" "GP" ...  
## $ sex : chr "F" "F" "F" "F" ...  
## $ age : int 18 17 15 15 16 16 16 17 15 15 ...  
## $ adress : chr "U" "U" "U" "U" ...  
## $ famsize : chr "GT3" "GT3" "LE3" "GT3" ...  
## $ Pstatus : chr "A" "T" "T" "T" ...  
## $ Medu : int 4 1 1 4 3 4 2 4 3 3 ...  
## $ Fedu : int 4 1 1 2 3 3 2 4 2 4 ...  
## $ Mjob : chr "at\_home" "at\_home" "at\_home" "health" ...  
## $ Fjob : chr "teacher" "other" "other" "services" ...  
## $ reason : chr "course" "course" "other" "home" ...  
## $ guardian : chr "mother" "father" "mother" "mother" ...  
## $ traveltime: int 2 1 1 1 1 1 1 2 1 1 ...  
## $ studytime : int 2 2 2 3 2 2 2 2 2 2 ...  
## $ failures : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ schoolsup : chr "yes" "no" "yes" "no" ...  
## $ famsup : chr "no" "yes" "no" "yes" ...  
## $ paid : chr "no" "no" "no" "no" ...  
## $ activities: chr "no" "no" "no" "yes" ...  
## $ nursery : chr "yes" "no" "yes" "yes" ...  
## $ higher : chr "yes" "yes" "yes" "yes" ...  
## $ internet : chr "no" "yes" "yes" "yes" ...  
## $ romantic : chr "no" "no" "no" "yes" ...  
## $ famrel : int 4 5 4 3 4 5 4 4 4 5 ...  
## $ freetime : int 3 3 3 2 3 4 4 1 2 5 ...  
## $ goout : int 4 3 2 2 2 2 4 4 2 1 ...  
## $ Dalc : int 1 1 2 1 1 1 1 1 1 1 ...  
## $ Walc : int 1 1 3 1 2 2 1 1 1 1 ...  
## $ health : int 3 3 3 5 5 5 3 1 1 5 ...  
## $ absences : int 4 2 6 0 0 6 0 2 0 0 ...  
## $ G1 : int 0 9 12 14 11 12 13 10 15 12 ...  
## $ G2 : int 11 11 13 14 13 12 12 13 16 12 ...  
## $ G3 : int 11 11 12 14 13 13 13 13 17 13 ...

## Las primeras filas de los datos

head(alcohol)

## school sex age adress famsize Pstatus Medu Fedu Mjob Fjob reason  
## 1 GP F 18 U GT3 A 4 4 at\_home teacher course  
## 2 GP F 17 U GT3 T 1 1 at\_home other course  
## 3 GP F 15 U LE3 T 1 1 at\_home other other  
## 4 GP F 15 U GT3 T 4 2 health services home  
## 5 GP F 16 U GT3 T 3 3 other other home  
## 6 GP M 16 U LE3 T 4 3 services other reputation  
## guardian traveltime studytime failures schoolsup famsup paid activities  
## 1 mother 2 2 0 yes no no no  
## 2 father 1 2 0 no yes no no  
## 3 mother 1 2 0 yes no no no  
## 4 mother 1 3 0 no yes no yes  
## 5 father 1 2 0 no yes no no  
## 6 mother 1 2 0 no yes no yes  
## nursery higher internet romantic famrel freetime goout Dalc Walc health  
## 1 yes yes no no 4 3 4 1 1 3  
## 2 no yes yes no 5 3 3 1 1 3  
## 3 yes yes yes no 4 3 2 2 3 3  
## 4 yes yes yes yes 3 2 2 1 1 5  
## 5 yes yes no no 4 3 2 1 2 5  
## 6 yes yes yes no 5 4 2 1 2 5  
## absences G1 G2 G3  
## 1 4 0 11 11  
## 2 2 9 11 11  
## 3 6 12 13 12  
## 4 0 14 14 14  
## 5 0 11 13 13  
## 6 6 12 12 13

## Obtener un resumen estadistico de cada variable

summary(alcohol)

## school sex age adress   
## Length:649 Length:649 Min. :15.00 Length:649   
## Class :character Class :character 1st Qu.:16.00 Class :character   
## Mode :character Mode :character Median :17.00 Mode :character   
## Mean :16.74   
## 3rd Qu.:18.00   
## Max. :22.00   
## famsize Pstatus Medu Fedu   
## Length:649 Length:649 Min. :0.000 Min. :0.000   
## Class :character Class :character 1st Qu.:2.000 1st Qu.:1.000   
## Mode :character Mode :character Median :2.000 Median :2.000   
## Mean :2.515 Mean :2.307   
## 3rd Qu.:4.000 3rd Qu.:3.000   
## Max. :4.000 Max. :4.000   
## Mjob Fjob reason guardian   
## Length:649 Length:649 Length:649 Length:649   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## traveltime studytime failures schoolsup   
## Min. :1.000 Min. :1.000 Min. :0.0000 Length:649   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.0000 Class :character   
## Median :1.000 Median :2.000 Median :0.0000 Mode :character   
## Mean :1.569 Mean :1.931 Mean :0.2219   
## 3rd Qu.:2.000 3rd Qu.:2.000 3rd Qu.:0.0000   
## Max. :4.000 Max. :4.000 Max. :3.0000   
## famsup paid activities nursery   
## Length:649 Length:649 Length:649 Length:649   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## higher internet romantic famrel   
## Length:649 Length:649 Length:649 Min. :1.000   
## Class :character Class :character Class :character 1st Qu.:4.000   
## Mode :character Mode :character Mode :character Median :4.000   
## Mean :3.931   
## 3rd Qu.:5.000   
## Max. :5.000   
## freetime goout Dalc Walc health   
## Min. :1.00 Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000   
## 1st Qu.:3.00 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:2.000   
## Median :3.00 Median :3.000 Median :1.000 Median :2.00 Median :4.000   
## Mean :3.18 Mean :3.185 Mean :1.502 Mean :2.28 Mean :3.536   
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:2.000 3rd Qu.:3.00 3rd Qu.:5.000   
## Max. :5.00 Max. :5.000 Max. :5.000 Max. :5.00 Max. :5.000   
## absences G1 G2 G3   
## Min. : 0.000 Min. : 0.0 Min. : 0.00 Min. : 0.00   
## 1st Qu.: 0.000 1st Qu.:10.0 1st Qu.:10.00 1st Qu.:10.00   
## Median : 2.000 Median :11.0 Median :11.00 Median :12.00   
## Mean : 3.659 Mean :11.4 Mean :11.57 Mean :11.91   
## 3rd Qu.: 6.000 3rd Qu.:13.0 3rd Qu.:13.00 3rd Qu.:14.00   
## Max. :32.000 Max. :19.0 Max. :19.00 Max. :19.00

## Analisis individual de cada una de las variables que conforman el set de datos

summary(alcohol$school)

## Length Class Mode   
## 649 character character

summary(alcohol$sex)

## Length Class Mode   
## 649 character character

summary(alcohol$age)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 15.00 16.00 17.00 16.74 18.00 22.00

summary(alcohol$adress)

## Length Class Mode   
## 649 character character

summary(alcohol$famsize)

## Length Class Mode   
## 649 character character

summary(alcohol$Pstatus)

## Length Class Mode   
## 649 character character

summary(alcohol$Medu)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 2.000 2.000 2.515 4.000 4.000

summary(alcohol$Fedu)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 1.000 2.000 2.307 3.000 4.000

summary(alcohol$Mjob)

## Length Class Mode   
## 649 character character

summary(alcohol$Fjob)

## Length Class Mode   
## 649 character character

summary(alcohol$reason)

## Length Class Mode   
## 649 character character

summary(alcohol$guardian)

## Length Class Mode   
## 649 character character

summary(alcohol$traveltime)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 1.000 1.000 1.569 2.000 4.000

summary(alcohol$studytime)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 1.000 2.000 1.931 2.000 4.000

summary(alcohol$failures)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0000 0.0000 0.0000 0.2219 0.0000 3.0000

summary(alcohol$schoolsup)

## Length Class Mode   
## 649 character character

summary(alcohol$famsup)

## Length Class Mode   
## 649 character character

summary(alcohol$paid)

## Length Class Mode   
## 649 character character

summary(alcohol$activities)

## Length Class Mode   
## 649 character character

summary(alcohol$nursery)

## Length Class Mode   
## 649 character character

summary(alcohol$higher)

## Length Class Mode   
## 649 character character

summary(alcohol$internet)

## Length Class Mode   
## 649 character character

summary(alcohol$romantic)

## Length Class Mode   
## 649 character character

summary(alcohol$famrel)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 4.000 4.000 3.931 5.000 5.000

summary(alcohol$freetime)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.00 3.00 3.00 3.18 4.00 5.00

summary(alcohol$goout)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 2.000 3.000 3.185 4.000 5.000

summary(alcohol$Dalc)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 1.000 1.000 1.502 2.000 5.000

summary(alcohol$Walc)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.00 1.00 2.00 2.28 3.00 5.00

summary(alcohol$health)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 2.000 4.000 3.536 5.000 5.000

summary(alcohol$absences)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 0.000 2.000 3.659 6.000 32.000

summary(alcohol$G1)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0 10.0 11.0 11.4 13.0 19.0

summary(alcohol$G2)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 10.00 11.00 11.57 13.00 19.00

summary(alcohol$G3)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 10.00 12.00 11.91 14.00 19.00