Proyecto Final

Elman Estuardo Camó Morales&Ever Geovany Sical López

2/6/2021

library(readr)  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.0.5

library(lubridate)

## Warning: package 'lubridate' was built under R version 4.0.5

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':  
##   
## date, intersect, setdiff, union

library(forecast)

## Warning: package 'forecast' was built under R version 4.0.5

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

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.0.5

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v tibble 3.1.1 v dplyr 1.0.5  
## v tidyr 1.1.3 v stringr 1.4.0  
## v purrr 0.3.4 v forcats 0.5.1

## Warning: package 'tibble' was built under R version 4.0.5

## Warning: package 'tidyr' was built under R version 4.0.5

## Warning: package 'purrr' was built under R version 4.0.5

## Warning: package 'dplyr' was built under R version 4.0.5

## Warning: package 'stringr' was built under R version 4.0.5

## Warning: package 'forcats' was built under R version 4.0.5

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x lubridate::as.difftime() masks base::as.difftime()  
## x lubridate::date() masks base::date()  
## x dplyr::filter() masks stats::filter()  
## x lubridate::intersect() masks base::intersect()  
## x dplyr::lag() masks stats::lag()  
## x lubridate::setdiff() masks base::setdiff()  
## x lubridate::union() masks base::union()

library(nortest)

csgo\_round\_snapshots <- read.csv("C:/Users/MasterYny/Desktop/csgo\_round\_snapshots.csv", header=TRUE)  
bitcoin<- read.csv("C:/Users/MasterYny/Desktop/final\_Bitcoin\_dataset.csv", header=TRUE)  
test<- read.csv("C:/Users/MasterYny/Desktop/test.csv", header=TRUE)

# Analisis 1

Demostrar si existe alguna correlacion entre las variables ct\_money y ct\_score

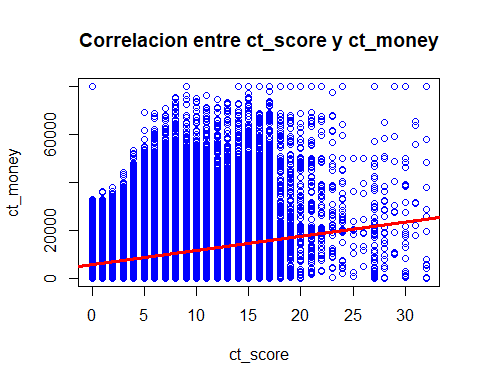
h0: No existe correlacion ente ambas variables

h1: Existe correlacion entre ambas variables

MODELO: T.CORRELACIOON

GRAFICO: DISPERCION

plot(csgo\_round\_snapshots$ct\_score,csgo\_round\_snapshots$ct\_money,col="blue",main="Correlacion entre ct\_score y ct\_money",xlab="ct\_score",ylab="ct\_money")  
abline(lm(csgo\_round\_snapshots$ct\_money ~ csgo\_round\_snapshots$ct\_score), col = 'red', lwd = 3)



cor(csgo\_round\_snapshots$ct\_score,csgo\_round\_snapshots$ct\_money)

## [1] 0.249416

0.249416<0.05

## [1] FALSE

Dado el analisis, la grafica y nuestro pvalue se puede determinar que la variable ct\_money y t\_money no es dependiente del ct\_score y t\_score por lo tanto “se afirma H1” Una investigacion que se realizo por cuenta propia quedo demostrado que en el juego de Counter Strike Gloval Ofencive el score no influye de forma directa en el dinero que se gana por ronda es decir no importa cuantas kills se tengan si se pierde la ronda se ganara menos dinero. (CLIP DE VIDO DEMOSTRATIVO)

# Análisis 2

Realize un analisis entre las variables bomb\_planted y map.

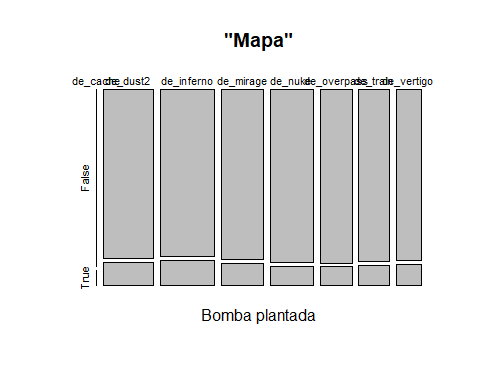
H0: La difuclta para plantar la bomba es la misma en todos los mapas

H1: La difuclta para plantar la bomba no es la misma en todos los mapas

MODELO:chi.test

GRAFICO: mosaico

tabla <- table(csgo\_round\_snapshots$map,csgo\_round\_snapshots$bomb\_planted, dnn=c("Bomba plantada"))  
mosaicplot(tabla, main = deparse("Mapa"))



tabla

## NA  
## Bomba plantada False True  
## de\_cache 134 11  
## de\_dust2 19474 2670  
## de\_inferno 20729 3082  
## de\_mirage 16483 2093  
## de\_nuke 17173 1852  
## de\_overpass 12742 1339  
## de\_train 12077 1414  
## de\_vertigo 9914 1223

chisq.test(tabla)

##   
## Pearson's Chi-squared test  
##   
## data: tabla  
## X-squared = 180.48, df = 7, p-value < 2.2e-16

2.2e-16<0.05

## [1] TRUE

Demostrando de forma numerico y grafica, el analisis nos indica que la dificultad para plantar la bomba es distinta en cada mapa por lo tanto “se afirma H1”

Este análisi puede tomar mejor sentido y forma si tomamos jugadas de jugadores profecionales e incluso en partidas casuales (CLIP DE VIDEO CON TOMAS PANORAMICAS Y SIMULACION)

# Analisis 3

Realize un analisis para demostrar cual es la media de kit de defuse en base al mapa

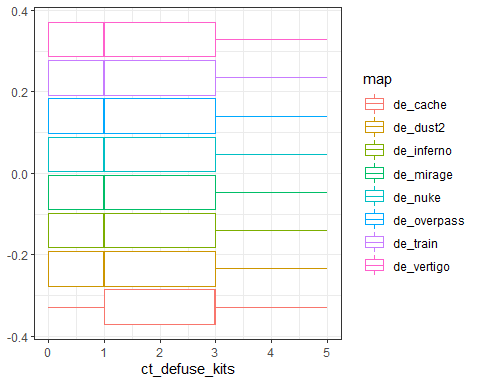
H0: No existe diferencia en la media de compra de kits de defuse por mapa

H1: En algunos casos existe difetencia en la media de compra de kits de defuse por mapa

Grafico: ggplot

Modelo: Anova

ggplot(data= csgo\_round\_snapshots, aes(x=ct\_defuse\_kits, y=,color=map ))+geom\_boxplot()+theme\_bw()



anova <- aov(csgo\_round\_snapshots$ct\_defuse\_kits~csgo\_round\_snapshots$map)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## csgo\_round\_snapshots$map 7 767 109.50 42.57 <2e-16 \*\*\*  
## Residuals 122402 314869 2.57   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dado el resultado podemos demostrar que la media de kit de defuse es de 1 por ronda esto explicado mejor en el (clip) por lo tanto “no se rechaza H0”

# Analisis 4

Determinar si el chaleco restante ct esta relacionada con el casco restante ct

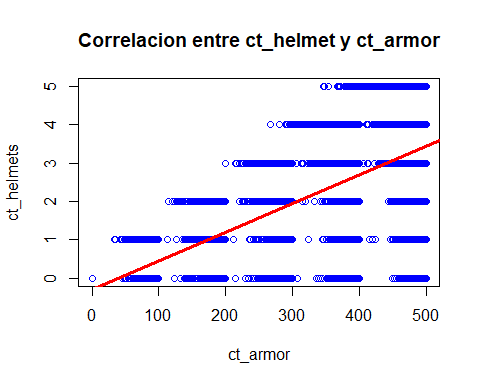
H0:La variable ct\_helmet no es dependiente de la variable ct\_armor

H1:La variable ct\_helmet es dependiente de la variable ct\_armor

Grafico: Dispersion

Modelo: Correlacion

plot(csgo\_round\_snapshots$ct\_armor,csgo\_round\_snapshots$ct\_helmets,col="blue",main="Correlacion entre ct\_helmet y ct\_armor",xlab="ct\_armor",ylab="ct\_helmets")  
abline(lm(csgo\_round\_snapshots$ct\_helmets ~ csgo\_round\_snapshots$ct\_armor), col = 'red', lwd = 3)



cor(csgo\_round\_snapshots$ct\_armor,csgo\_round\_snapshots$ct\_helmets)

## [1] 0.6945763

Dada la grafica y el analisis podemos demostrar que el helmet es dependiente de armor por tanto no se rechaza H0 (CLIP DEMOSTRATIVO)

# Analisis 5

Determine si existe diferencia de media en compras de AK47 segun el mapa jugado

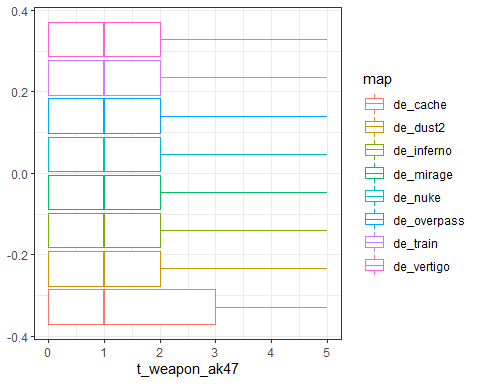
Grafica:ggplot

Modelo: anova

h0: no existe diferencia en la compra de 1 ak47

h1: en algunos casos existe diferenica en la compra de 1 ak47

ggplot(data= csgo\_round\_snapshots, aes(x=t\_weapon\_ak47, y=,color=map ))+geom\_boxplot()+theme\_bw()



anova <- aov(csgo\_round\_snapshots$t\_weapon\_ak47~csgo\_round\_snapshots$map)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## csgo\_round\_snapshots$map 7 803 114.75 61.9 <2e-16 \*\*\*  
## Residuals 122402 226900 1.85   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dado el resultado y la grafica se puede determinar que la media sera de por lo menos 1 ak47 comprada en las rodondas por lo tanto “no se rechaza H0”

# Analisis 6

Detenerminar si el tiempo restante tendra una diferencia significativa en el estado en el que se enconbtro la bomba antes de acabar la ronda

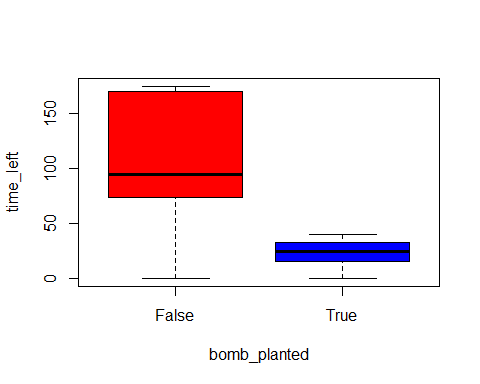
H0: El tiempo restante tendra una diferencia significativa en el estado en el que se enconbtro la bomba antes de acabar la ronda

H1: El tiempo restante no tendra una diferencia significativa en el estado en el que se enconbtro la bomba antes de acabar la ronda

Grafica: Boxplot

Modelo: T.test

boxplot(time\_left~bomb\_planted, data=csgo\_round\_snapshots, col=c("red","blue"))



t.test(csgo\_round\_snapshots$time\_left~csgo\_round\_snapshots$bomb\_planted, alternative="two.sided")

##   
## Welch Two Sample t-test  
##   
## data: csgo\_round\_snapshots$time\_left by csgo\_round\_snapshots$bomb\_planted  
## t = 469.86, df = 97646, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 83.6614 84.3623  
## sample estimates:  
## mean in group False mean in group True   
## 107.27846 23.26661

Dado el analisis y la grafica se puede demostrar que H0 no se rechaza y se puede demostrar de forma real con un clip que el tiempo restante esta relacionado con la bomba plantada

# Analisis 7

Determine si existe diferencia de media en compras de m4a4 ct segun el mapa jugado

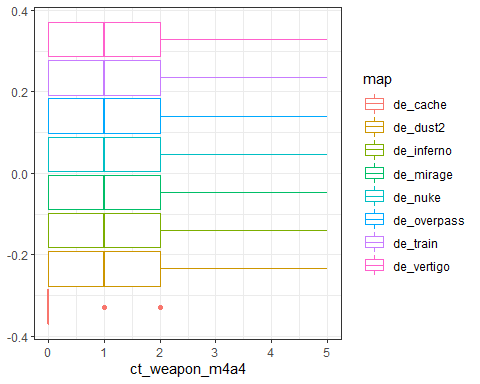
Grafica:ggplot

Modelo: anova

h0: no existe diferencia en la compra de 1 m4a4

h1: en algunos casos existe diferenica en la compra de 1 m4a4

ggplot(data= csgo\_round\_snapshots, aes(x=ct\_weapon\_m4a4, y=,color=map ))+geom\_boxplot()+theme\_bw()



anova <- aov(csgo\_round\_snapshots$ct\_weapon\_m4a4~csgo\_round\_snapshots$map)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## csgo\_round\_snapshots$map 7 487 69.52 46.9 <2e-16 \*\*\*  
## Residuals 122402 181442 1.48   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dado el resultado y la grafica se puede determinar que la media sera de por lo menos 1 m4a4 CT comprada en las rodondas por lo tanto “no se rechaza H0”

# DataSet de Datos Aviones

# Analisis 8

Determinar si la edad del individuo es a partir de la distancia que viaja cada uno de estos demuestrelo graficamente como analiticamente

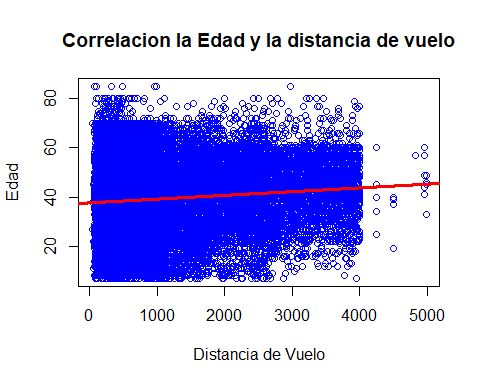
Modelo = Correlacion

Grafico = Dispersion

H0: La edad de la persona no influye en la distancia de vuelo

H1: La edad de la persona influye en la distancia de vuelo

plot(test$Flight.Distance,test$Age,col="blue",main="Correlacion la Edad y la distancia de vuelo",xlab="Distancia de Vuelo",ylab="Edad")  
abline(lm(test$Age ~ test$Flight.Distance), col = 'red', lwd = 3)



cor(test$Flight.Distance,test$Age)

## [1] 0.09940871

#pvalue<0.05  
0.09940871<0.05

## [1] FALSE

Como se demuestra que el pvalue es menor no tenemos pruebas suficientes para rechazar H0 y por ende podemos confirmar de que la edad de la persona no influye en la distancia de vuelo que viaja el individuo

# Analisis 9

Determine graficamente y por medio de un analisis si la clase en la que vuela el individuo esta relacionado en base al tipo de cliente que es el individuo.

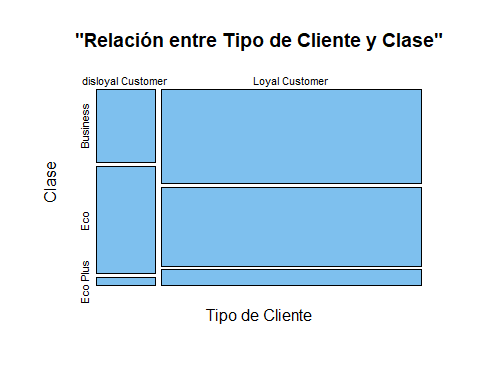
Modelo = Test chi cuadrado

Grafica = Mosaic plot

H0: A mejor tipo de cliente no es mejor es la clase en la que vuela

H1: A mejor tipo de cliente mejor la clase en la que vuela

tablaclc <- table(test$Customer.Type,test$Class, dnn=c("Tipo de Cliente","Clase"))  
mosaicplot(tablaclc,color = "skyblue2", main = deparse("Relación entre Tipo de Cliente y Clase"))



tablaclc

## Clase  
## Tipo de Cliente Business Eco Eco Plus  
## disloyal Customer 1875 2724 200  
## Loyal Customer 10620 8840 1717

chisq.test(tablaclc)

##   
## Pearson's Chi-squared test  
##   
## data: tablaclc  
## X-squared = 380.35, df = 2, p-value < 2.2e-16

#pvalue<0.05  
2.2e-16<0.05

## [1] TRUE

Existen suficientes pruebas para afirmar H1 por ende se puede decir que la lealtad del cliente si influye en el tipo de clase que este viaja, graficamente observamos que los tipos de cliente “Leales” tienden a viajar mas en primera clase de aque aquellos que “No son leales” en la grafica se muestra que este tipo de cliente tienede a viajar mas economicamente.

# Analisis 10

Por medio de un analisis comprobar graficamente y analiticamente si existe una diferencia significativa entre la califiacion de la comida y bebida que ponen los individuos en base al genero sea hombre o mujer.

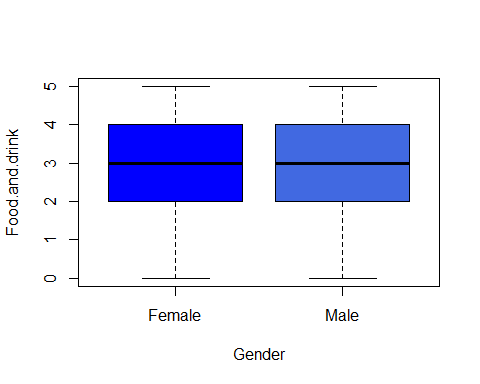
Modelo = t.test

Grafica = Boxplot

H0:La satisfaccion de comida y bebida no depende de la genero

H1:La satisfaccion de comida y bebida depende del genero

boxplot(Food.and.drink~Gender, data=test, col=c("blue","royalblue"))



t.test(test$Food.and.drink~test$Gender, alternative="two.sided")

##   
## Welch Two Sample t-test  
##   
## data: test$Food.and.drink by test$Gender  
## t = 2.3246, df = 25954, p-value = 0.0201  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.006022549 0.070794567  
## sample estimates:  
## mean in group Female mean in group Male   
## 3.234285 3.195876

0.0201<0.05

## [1] TRUE

Podemos ver que nuestro pvalue es menor asi afirmando H1 diciendo que la nota que asignan los individuos depende del genero, aunque graficamente no se ve podemos decir que las mujeres califican un poco mas alto la comida y bebida, esto se ve en la media de cada uno, la diferencia no es tan grande pero nos muestra que existe como tal.

# Analisis 11

La calificacion que dan los individuos del nivel de abordaje estara relacionado co la calificacion que se le da al nivel de entretenimiento que se le da en el vuelo.

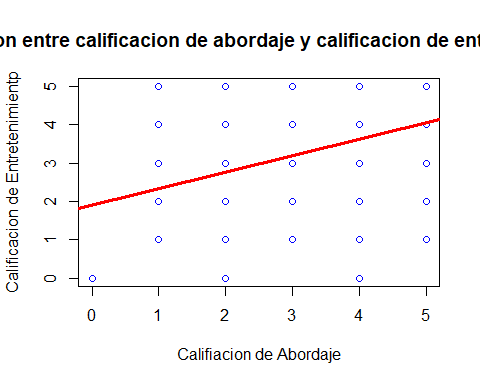
Modelo = Correlacion

Grafico = Dispersion

H0: La nota asiganada nivel de abordaje no depende de la nota que se asigna al entretenimiento del vuelo

H1: La nota asiganada nivel de abordaje depende de la nota que se asigna al entretenimiento del vuelo

plot(test$On.board.service,test$Inflight.entertainment,col="blue",main="Correlaion entre calificacion de abordaje y calificacion de entretenimiento",xlab="Califiacion de Abordaje",ylab="Calificacion de Entretenimientp")  
abline(lm(test$Inflight.entertainment ~ test$On.board.service), col = 'red', lwd = 3)



cor(test$On.board.service,test$Inflight.entertainment)

## [1] 0.4122661

#pvalue<0.05  
0.4122661<0.05

## [1] FALSE

No existen suficientes pruebas para rechazar H0 asi que podeemos decir que el nivel de abordaje que se brinda no tiene depende de la nota asignada al entretenimiento.

# Analisis 12

Realice un analisis donde se demuestre si el tipo de satisfaccion que asigno el individuo depende si este es hombre o mujer

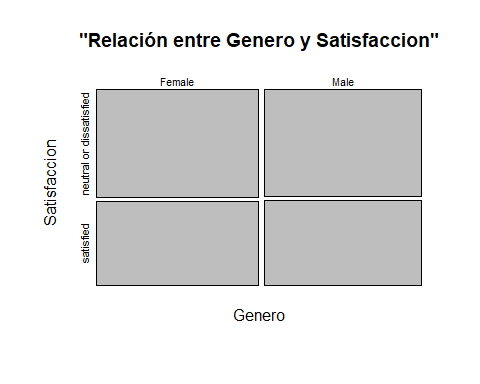
Modelo = Test chi cuadrado

Grafica = Mosaic plot

H0: La nota asignada no depende si es hombre o mujer

H1: La nota asignada depende si es hombre o mujer

tablagen <- table(test$Gender,test$satisfaction, dnn=c("Genero","Satisfaccion"))  
mosaicplot(tablagen, main = deparse("Relación entre Genero y Satisfaccion"))



tablagen

## Satisfaccion  
## Genero neutral or dissatisfied satisfied  
## Female 7437 5735  
## Male 7136 5668

chisq.test(tablagen)

##   
## Pearson's Chi-squared test with Yates' continuity correction  
##   
## data: tablagen  
## X-squared = 1.3682, df = 1, p-value = 0.2421

#pvalue<0.05  
0.2421<0.05

## [1] FALSE

El pvalue obtenido es mayor asi que no se rechaza H0, y podemos decir que el nivel de satisfaccion que recibe cada uno no depende si este sea hombre o mujer.

# Analisis 13

Realice un analisis para comprobarar si la satisfaccion que asigna cada individuo depende de la edad del individuo

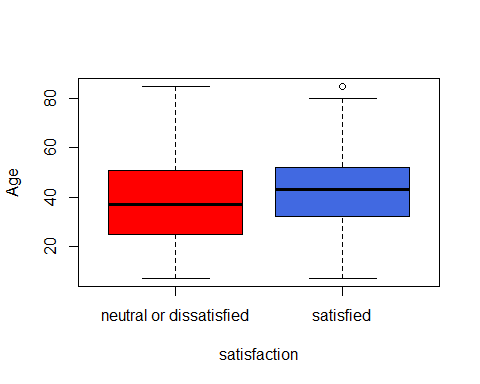
Modelo = t.test

Grafica = Boxplot

H0: La edad no influye en el nivel de satisfaccion

H1: La edad influye en el nivel de satisfaccion

boxplot(Age~satisfaction, data=test, col=c("red","royalblue"))



t.test(test$Age~test$satisfaction, alternative="two.sided")

##   
## Welch Two Sample t-test  
##   
## data: test$Age by test$satisfaction  
## t = -20.346, df = 25974, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -4.069169 -3.354047  
## sample estimates:  
## mean in group neutral or dissatisfied mean in group satisfied   
## 37.99163 41.70324

2.2e-16<0.05

## [1] TRUE

Debido a los suficientes datos que tenemos, afirmamos H1, y llegamos a la conclusion La edad influye en el nivel de satisfaccion que asignan cada indiduo llega a ser mayor en personas que tienen una edad de 41 años.

# Analisis 14

Comprobar si el retraso de salida del avion esta relacionado en base al retraso de llegada de este mismo

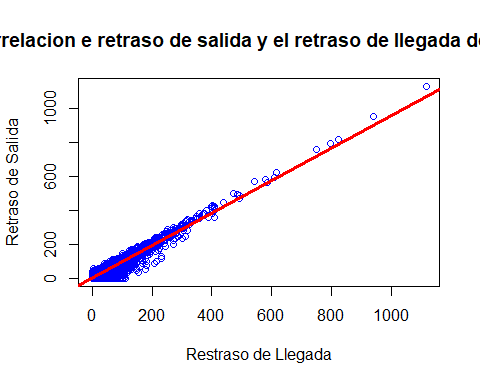
Modelo = Correlacion

Grafico = Dispersion

H0: El retraso del llegada del avion no depende del retraso de salida del mismo

H1: El retraso del llegada del avion depende del retraso de salida del mismo

plot(test$Arrival.Delay.in.Minutes,test$Departure.Delay.in.Minutes,col="blue",main="Correlacion e retraso de salida y el retraso de llegada del avion",xlab="Restraso de Llegada",ylab="Retraso de Salida")  
abline(lm(test$Departure.Delay.in.Minutes ~ test$Arrival.Delay.in.Minutes), col = 'red', lwd = 3)



cor(test$Arrival.Delay.in.Minutes,test$Departure.Delay.in.Minutes)

## [1] NA

#pvalue<0.05  
NA<0.05

## [1] NA

Debido a que el grado de correlacion es NA podemos determinar que no existe una relacion directa sobre el retraso de salida y el retraso de entrada que tiene el avion asi que no podemos rechazar H0.

# Set de datos Cryptomonedas

# Analisis 15

Realice un Analsis para ver el movimento de la tasa del bitcoin de aca a 15 meses en el periodo de 2021 encontrar cual es la prediccion con menos tasa de error.

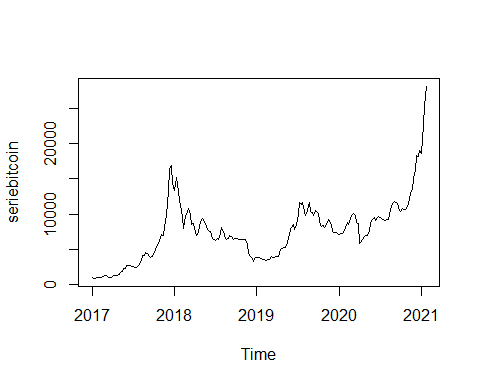
Modelo: Window, ARIMA, HoltWinters

Grafico: Series de Tiempo

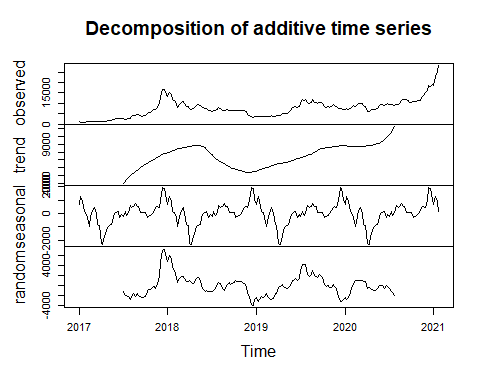
bitcoin$fecha = as.Date(bitcoin$Timestamp,format="%Y-%m-%d","%H:%M:%S")  
  
bitcoin$sem = week(bitcoin$fecha)  
  
bitcoin$year = year(bitcoin$fecha)  
  
bitcoinsemanal = bitcoin %>% group\_by(year,sem) %>% summarise(media=mean(High))

## `summarise()` has grouped output by 'year'. You can override using the `.groups` argument.

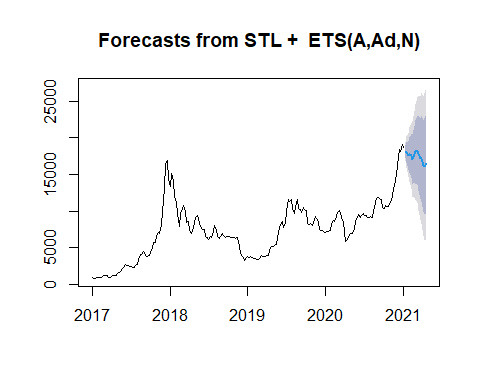
seriebitcoin = ts(bitcoinsemanal$media, start=c(2017,1),frequency = 52)  
  
plot(seriebitcoin)



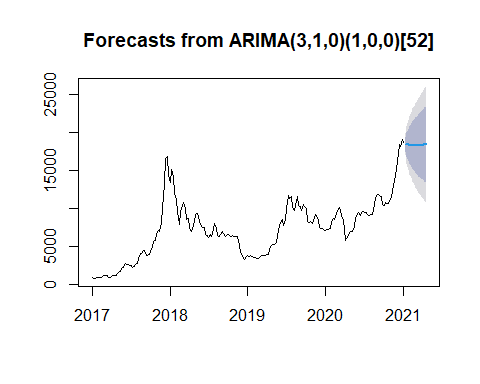
componentes = decompose(seriebitcoin)  
  
plot(componentes)



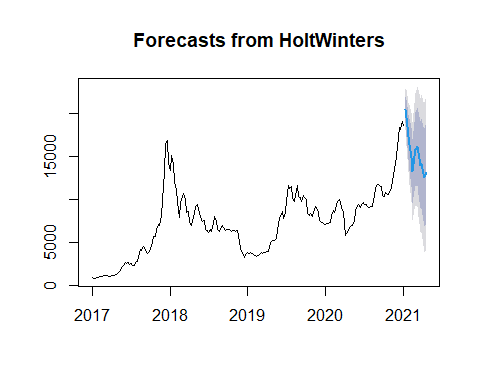
Bitcoin2021 = window(seriebitcoin, start=2017,end=2021)  
  
pre1=forecast(Bitcoin2021,15)  
  
plot(pre1)



ms=auto.arima(Bitcoin2021)  
  
pre2=forecast(ms,15)  
  
plot(pre2)



modHT = HoltWinters(Bitcoin2021)  
  
pre3 = forecast(modHT,15)  
  
plot(pre3)



accuracy(pre1)

## ME RMSE MAE MPE MAPE MASE ACF1  
## Training set 32.77384 638.0207 472.1486 0.9344257 9.06548 0.1055693 0.1305464

accuracy(pre2)

## ME RMSE MAE MPE MAPE MASE  
## Training set 59.29593 708.3423 470.0773 0.7730931 6.377754 0.1051062  
## ACF1  
## Training set -0.003173689

accuracy(pre3)

## ME RMSE MAE MPE MAPE MASE  
## Training set -29.23782 1186.309 800.1833 -0.7790805 12.21983 0.1789157  
## ACF1  
## Training set 0.4148837

Llegamos a la conlcusion de que el metodo con menos tasa de error es el Arima asi concluyendo que la subida del bitcon permanecera estable en un aproximado de 190,000 dolares

# Analisis 16

Demuestre medio un histograma si existe diferencia entre las medias de subida y bajada ademas de comprobar si existe normalidad en los datos

Modelo: Test de Normalidad

Grafico: Histograma

H0: Existe Normalidad en los datos

H1: No existe normalidad en los datos

summary(bitcoin)

## X Timestamp Open High   
## Min. :2625377 Length:2102400 Min. : 752 Min. : 752.8   
## 1st Qu.:3150977 Class :character 1st Qu.: 4083 1st Qu.: 4084.4   
## Median :3676576 Mode :character Median : 7350 Median : 7354.6   
## Mean :3676576 Mean : 7463 Mean : 7467.8   
## 3rd Qu.:4202176 3rd Qu.: 9637 3rd Qu.: 9643.1   
## Max. :4727776 Max. :28997 Max. :29010.3   
## Low Close Volume\_.BTC. Volume\_.Currency.   
## Min. : 751.3 Min. : 752 Min. : 0.0000 Min. : 0   
## 1st Qu.: 4080.3 1st Qu.: 4083 1st Qu.: 0.3737 1st Qu.: 2056   
## Median : 7346.0 Median : 7350 Median : 1.9146 Median : 11618   
## Mean : 7457.7 Mean : 7463 Mean : 7.0898 Mean : 52031   
## 3rd Qu.: 9630.6 3rd Qu.: 9637 3rd Qu.: 6.6409 3rd Qu.: 44380   
## Max. :28963.7 Max. :28997 Max. :1616.0600 Max. :10445988   
## Weighted\_Price fecha sem year   
## Min. : 752 Min. :2017-01-01 Min. : 1.00 Min. :2017   
## 1st Qu.: 4083 1st Qu.:2018-01-01 1st Qu.:14.00 1st Qu.:2018   
## Median : 7350 Median :2019-01-01 Median :27.00 Median :2019   
## Mean : 7463 Mean :2018-12-31 Mean :26.57 Mean :2018   
## 3rd Qu.: 9637 3rd Qu.:2020-01-01 3rd Qu.:40.00 3rd Qu.:2020   
## Max. :28986 Max. :2020-12-31 Max. :53.00 Max. :2020

par(mfrow=c(1,2))  
  
hist(bitcoin$High,col="red")  
mediaus <- mean(bitcoin$High)  
abline(v=mediaus,col="black")  
mediaus

## [1] 7467.775

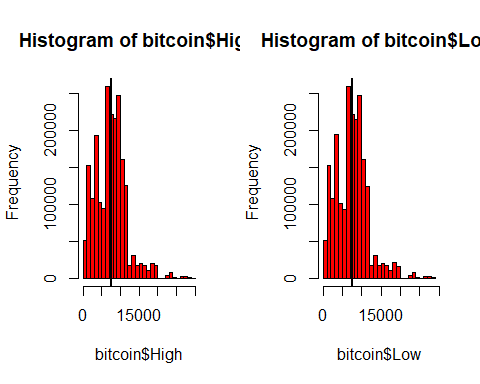
medianaus <- median(bitcoin$High)  
abline(v=medianaus,col="black")  
medianaus

## [1] 7354.56

hist(bitcoin$Low,col="red")  
medial <- mean(bitcoin$Low)  
abline(v=medial,col="black")  
medial

## [1] 7457.686

medianal <- median(bitcoin$Low)  
abline(v=medianal,col="black")



medianal

## [1] 7346

lillie.test(bitcoin$High)

##   
## Lilliefors (Kolmogorov-Smirnov) normality test  
##   
## data: bitcoin$High  
## D = 0.061058, p-value < 2.2e-16

lillie.test(bitcoin$Low)

##   
## Lilliefors (Kolmogorov-Smirnov) normality test  
##   
## data: bitcoin$Low  
## D = 0.060841, p-value < 2.2e-16

2.2e-16<0.05

## [1] TRUE

Existe diferencia estadistica para poder rechazar H0, por ende ninguna de las graficas muestra que tengan normalidad como tal.

# Analisis 17

Tendra alguna correlacion el Precio ponderado del Bitcon y el precio de cerrada que se le asigna tambien ha este.

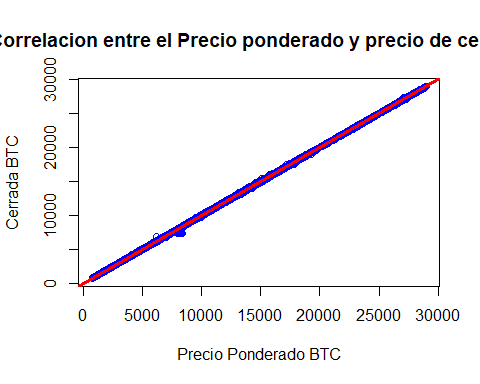
Modelo: Correlacion

Grafico: Dispersion

H0: Ambas variables tienen una relacion

H1: Ninguna de las variables tiene una relacion

plot(bitcoin$Weighted\_Price,bitcoin$Close,col="blue",main="Correlacion entre el Precio ponderado y precio de cerrada",xlab="Precio Ponderado BTC",ylab="Cerrada BTC")  
abline(lm(bitcoin$Close ~ bitcoin$Weighted\_Price), col = 'red', lwd = 3)



cor(bitcoin$Weighted\_Price,bitcoin$Close)

## [1] 0.9999964

0.9999964<0.05

## [1] FALSE

No tenemos suficientes datos para poder rechazar H0 asi que por ende la vairable de ponderado BTC tiene una relacion con el precio de cerrada que se le da al BTC.

# Analisis 18

Prufundizar en el mercado de divisas para observar los mas relevantes si ha aumentado o disminuido hasa el año actual

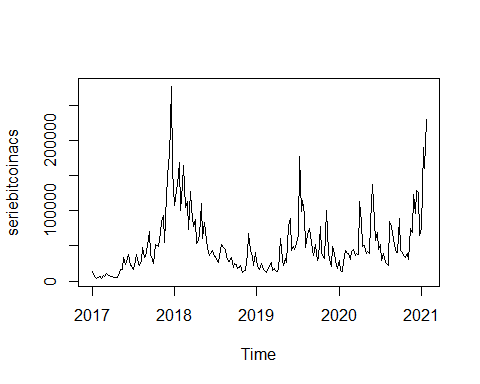
Modelo: Window, ARIMA, HoltWinters

Grafico: Series de Tiempo

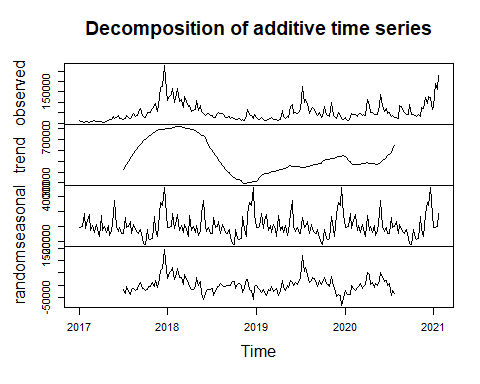
bitcoinacs = bitcoin %>% group\_by(year,sem)%>%summarise(media=mean(Volume\_.Currency.))

## `summarise()` has grouped output by 'year'. You can override using the `.groups` argument.

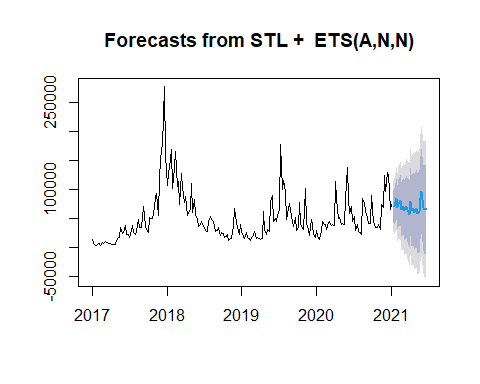
seriebitcoinacs = ts(bitcoinacs$media, start=c(2017,1),frequency = 52)  
  
plot(seriebitcoinacs)



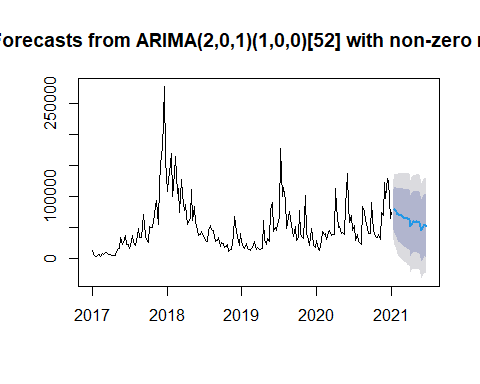
componentesacs = decompose(seriebitcoinacs)  
  
plot(componentesacs)



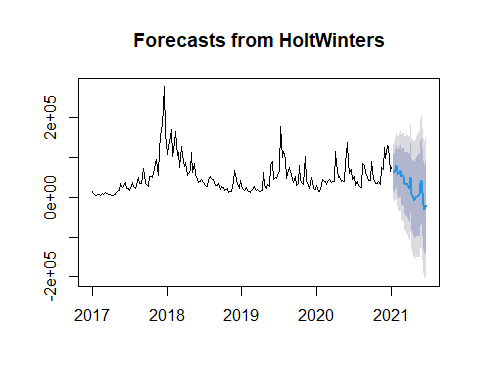
Bitcoin2021acs = window(seriebitcoinacs, start=2017,end=2021)  
  
pre1acs=forecast(Bitcoin2021acs,24)  
  
plot(pre1acs)



macs=auto.arima(Bitcoin2021acs)  
  
pre2acs=forecast(macs,24)  
  
plot(pre2acs)



modHTacs = HoltWinters(Bitcoin2021acs)  
  
pre3acs = forecast(modHTacs,24)  
  
plot(pre3acs)



accuracy(pre1acs)

## ME RMSE MAE MPE MAPE MASE ACF1  
## Training set 592.5956 21673.52 15609.22 -6.86533 43.63247 0.3264423 0.07046058

accuracy(pre2acs)

## ME RMSE MAE MPE MAPE MASE  
## Training set 321.2399 24494.06 16074.59 -22.60393 40.17604 0.3361749  
## ACF1  
## Training set 0.002468418

accuracy(pre3acs)

## ME RMSE MAE MPE MAPE MASE ACF1  
## Training set -1508.992 33976.46 23780.23 -16.15738 67.47059 0.4973261 0.1421443

# Analisis 19

Explicar cual fue el volumen de bitcoin, ademas de hacer una prediccon de como cambiara a 8 meses despues

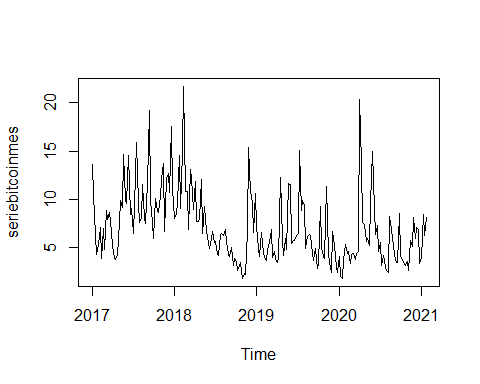
Modelo: Window, ARIMA, HoltWinters

Grafico: Series de Tiempo

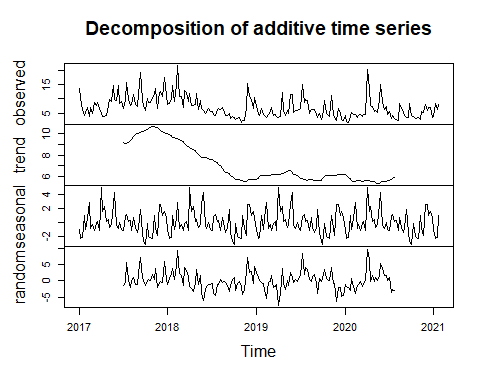
bitcoinmens = bitcoin %>% group\_by(year,sem) %>% summarise(media=mean(Volume\_.BTC.))

## `summarise()` has grouped output by 'year'. You can override using the `.groups` argument.

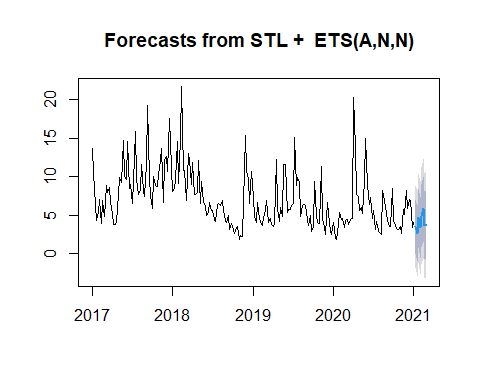
seriebitcoinmes = ts(bitcoinmens$media, start=c(2017,1),frequency = 52)  
  
plot(seriebitcoinmes)



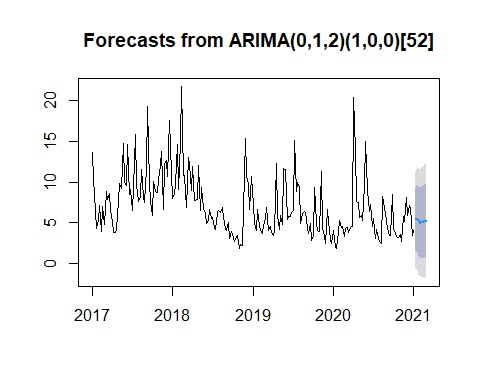
componentesmes = decompose(seriebitcoinmes)  
  
plot(componentesmes)



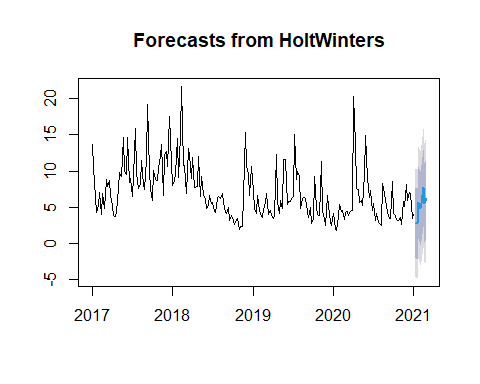
Bitcoin2021mes = window(seriebitcoinmes, start=2017,end=2021)  
  
pre1mes=forecast(Bitcoin2021mes,8)  
  
plot(pre1mes)



mse=auto.arima(Bitcoin2021mes)  
  
pre2mes=forecast(mse,8)  
  
plot(pre2mes)



modHTmes = HoltWinters(Bitcoin2021mes)  
  
pre3mes = forecast(modHTmes,8)  
  
plot(pre3mes)



accuracy(pre1mes)

## ME RMSE MAE MPE MAPE MASE  
## Training set -0.07479393 2.694079 1.967668 -11.92544 31.33103 0.4903657  
## ACF1  
## Training set 0.1793533

accuracy(pre2mes)

## ME RMSE MAE MPE MAPE MASE  
## Training set -0.145267 2.937541 2.132617 -15.61784 32.67439 0.5314729  
## ACF1  
## Training set 0.01882687

accuracy(pre3mes)

## ME RMSE MAE MPE MAPE MASE ACF1  
## Training set 0.01869063 3.74237 2.458904 -15.61297 47.90049 0.6127874 0.2218521

Podemos observar que la fecha mas relevante durante el periodo de finales de 2018 e inicios del 2019 debido a una abaja del precio en 2018

# Analisis 20

Analizar graficamente y analiticamente si existe una relacion entre Volumen del BTC y el Volumende de Divisas.

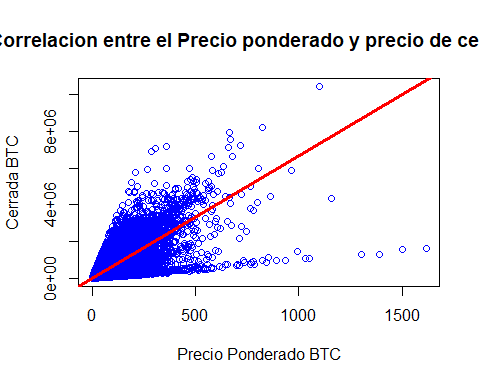
Modelo: Correlacion

Grafico: Dispersion

H0: Existe una relacion entra la variable BTC y el Volumen de Divisas

H1: No existe una relacion entra la variable BTC y el Volumen de Divisas

plot(bitcoin$Volume\_.BTC.,bitcoin$Volume\_.Currency.,col="blue",main="Correlacion entre el Precio ponderado y precio de cerrada",xlab="Precio Ponderado BTC",ylab="Cerrada BTC")  
abline(lm(bitcoin$Volume\_.Currency. ~ bitcoin$Volume\_.BTC.), col = 'red', lwd = 3)



cor(bitcoin$Volume\_.BTC.,bitcoin$Volume\_.Currency.)

## [1] 0.8418369

0.9999964<0.05

## [1] FALSE

No tenemos suficientes datos para afirmar lo H1 asi que no se rechaza H0 ya que el valor es demasiado alto y a su vez es positivo asi que podemos decir que existe una relacion entre ambas variables