ProyectoFinal

Mariana López y Jazmin Guzmán

1/6/2021

# Proyecto Final

lectura <- read.csv("HabitoLeer.csv")

Esto seria para leerlo al estar guardado en una carpeta, pero ya esta subido en github y aquí por esoo no lo hacemos así: lectura <- read.csv(“C:/Users/1/Desktop/Estadistica 2/Estadistica 2/Proyecto/HabitoLeer.csv”)

require(stats)  
library (corrplot)

## corrplot 0.88 loaded

library(ggplot2)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

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

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✓ tibble 3.1.2 ✓ purrr 0.3.4  
## ✓ tidyr 1.1.3 ✓ stringr 1.4.0  
## ✓ readr 1.4.0 ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(lubridate)

##   
## Attaching package: 'lubridate'

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

library(fpp2)

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

## ── Attaching packages ────────────────────────────────────────────── fpp2 2.4 ──

## ✓ forecast 8.15 ✓ expsmooth 2.3   
## ✓ fma 2.4

##

## 1

Explorando los datos

summary(lectura)

## Age Sex Race Marital.status.   
## Min. :16.00 Length:2832 Length:2832 Length:2832   
## 1st Qu.:32.00 Class :character Class :character Class :character   
## Median :48.00 Mode :character Mode :character Mode :character   
## Mean :47.28   
## 3rd Qu.:62.00   
## Max. :93.00   
## Education Employement Incomes   
## Length:2832 Length:2832 Length:2832   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
## How.many.books.did.you.read.during.last.12months.  
## Min. : 0.00   
## 1st Qu.: 2.75   
## Median : 6.00   
## Mean :16.68   
## 3rd Qu.:20.00   
## Max. :97.00   
## Read.any.printed.books.during.last.12months.  
## Length:2832   
## Class :character   
## Mode :character   
##   
##   
##   
## Read.any.audiobooks.during.last.12months.  
## Length:2832   
## Class :character   
## Mode :character   
##   
##   
##   
## Read.any.e.books.during.last.12months. Last.book.you.read..you.  
## Length:2832 Length:2832   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
## Do.you.happen.to.read.any.daily.news.or.newspapers.  
## Length:2832   
## Class :character   
## Mode :character   
##   
##   
##   
## Do.you.happen.to.read.any.magazines.or.journals.  
## Length:2832   
## Class :character   
## Mode :character   
##   
##   
##

### Tenemos distintas variables sin embargo los unicos enteros son la edad y la cantidad de libros leidos al año.

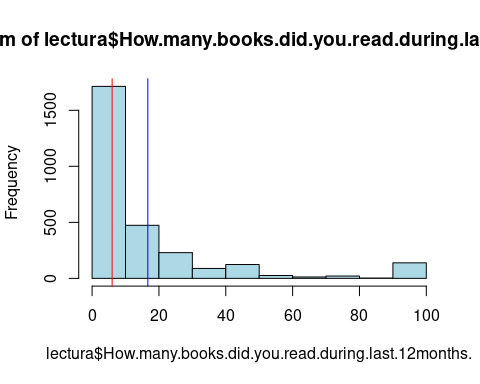
## 2

Promedio de libros al año

summary(lectura$How.many.books.did.you.read.during.last.12months.)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 2.75 6.00 16.68 20.00 97.00

hist(lectura$How.many.books.did.you.read.during.last.12months.,col="lightblue")  
medialib <- mean(lectura$How.many.books.did.you.read.during.last.12months.)  
  
abline(v=medialib,col="blue")  
  
medianlib <- median(lectura$How.many.books.did.you.read.during.last.12months.)  
abline(v=medianlib,col="red")

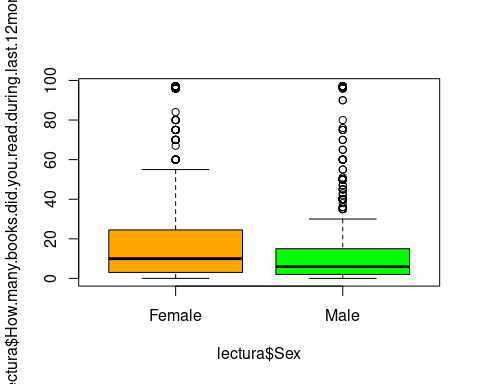
 ### Respuesta. La media de libros leidos por año es de 16.68 y una mediana de 6 (de rojo), siendo más confiable la mediana en este tipo de casos porque hay un sesgo hacia la derecha en los datos.

## Pregunta 3.

Si existe alguna diferencia entre hombres y mujeres y la cantidad de libros leidos durante todo el año.

H0: No hay diferencia, hombres = mujeres H1: Existe una diferencia, hombres =/= mujeres

boxplot(lectura$How.many.books.did.you.read.during.last.12months. ~ lectura$Sex, col=c("orange","green"))



t.test(lectura$How.many.books.did.you.read.during.last.12months.~lectura$Sex, alternative = "two.sided")

##   
## Welch Two Sample t-test  
##   
## data: lectura$How.many.books.did.you.read.during.last.12months. by lectura$Sex  
## t = 7.2566, df = 2734.5, p-value = 5.147e-13  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 4.610387 8.024509  
## sample estimates:  
## mean in group Female mean in group Male   
## 19.69439 13.37694

5.147e-13 < 0.05

## [1] TRUE

### Respuesta.

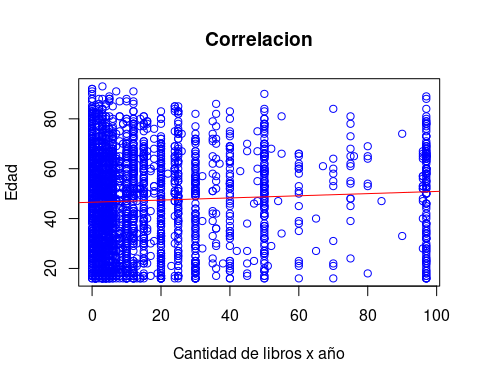
H1: Existe una diferencia, hombres =/= mujeres <https://elpais.com/cultura/2020/02/21/actualidad/1582288090_293986.html>

## pregunta 4

¿La edad influye en la cantidad de libros leidos durante todo el año?

Ho. no es relevante por que dependerá más de factores como el tiempo. H1. si es relevante por que en base la edad se va incrementando la cantidad de libros leídos.

plot(lectura$How.many.books.did.you.read.during.last.12months., lectura$Age,col="blue",main="Correlacion",xlab="Cantidad de libros x año",ylab="Edad")  
model <- lm(lectura$Age ~ lectura$How.many.books.did.you.read.during.last.12months.)  
abline(model,col="red")



cor(lectura$How.many.books.did.you.read.during.last.12months., lectura$Age)

## [1] 0.05496802

### Respuesta. No influye la edad de la persona en la cantidad de libros leidos al año.

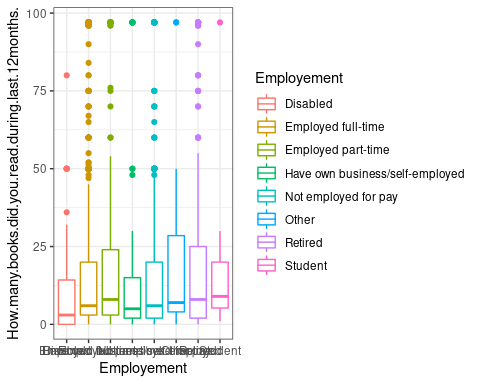
<https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2020/EstSociodemo/MOLEC2019_04.pdf>

## pregunta 5

¿Una persona con empleo lee más que una desempleada?

Ho. No por que depende mas del tiempo que la persona tenga. (no hay relacion empleado = desempleado) H1. Si por que por razones laborales puede incrementar su lista de libros leídos. (hay relacion empleado =/= desempleado)

ggplot(data = lectura, aes(x= Employement , y= How.many.books.did.you.read.during.last.12months., color= Employement))+ geom\_boxplot()+theme\_bw()



anova <- aov(lectura$How.many.books.did.you.read.during.last.12months.~lectura$Employement)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## lectura$Employement 7 15066 2152.3 3.881 0.000325 \*\*\*  
## Residuals 2824 1565976 554.5   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

0.000325 < 0

## [1] FALSE

### Respuesta. Si existe diferencia entre alguno de los grupos segun tengan empleo o no.

<https://pruebat.org/Inicio/ConSesion/Breves/verBreve/1404-por-que-cada-vez-hay-menos-lectores>

## pregunta 6

¿Los ingresos son factor importante en el consumo de libros?

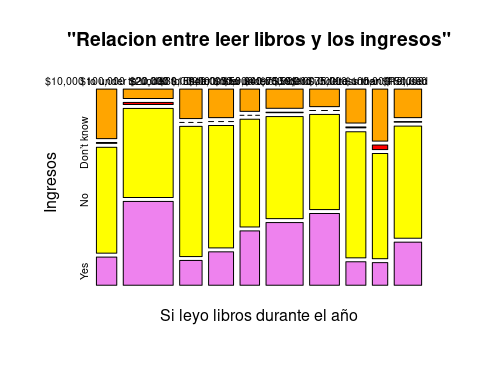
Ho. No ya que podemos consumir libros de manera digital o bien libros en otras fuentes. (variables independientes) H1. Si puesto que no se pueden adquirir todos los libros deseados sin tener ingresos suficientes. (variables dependientes)

tab <- table(lectura$Incomes, lectura$Read.any.e.books.during.last.12months., dnn = c("Si leyo libros durante el año","Ingresos"))  
 tab

## Ingresos  
## Si leyo libros durante el año Don’t know No Yes  
## $10,000 to under $20,000 58 1 124 33  
## $100,000 to under $150,000 27 6 256 241  
## $20,000 to under $30,000 38 0 168 32  
## $30,000 to under $40,000 41 0 176 48  
## $40,000 to under $50,000 25 0 121 61  
## $50,000 to under $75,000 41 1 218 134  
## $75,000 to under $100,000 30 0 163 123  
## 9$100,000 to under $150,000 39 1 145 27  
## Less than $10,000 46 4 93 20  
## Refused 45 1 177 68

### Se ve que quienes tienen un ingreso de 100,00 por debajo de 150,000 dijeron haber leido algun libro al año diferencia que se contrapone con los que ganan menos de 10,000. Por tanto, desde aquí nos damos cuenta que los ingresos si son factor importante.

mosaicplot(tab, main = deparse("Relacion entre leer libros y los ingresos"), color = c("orange","red","yellow","violet"))



chisq.test(tab)

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

##   
## Pearson's Chi-squared test  
##   
## data: tab  
## X-squared = 293.46, df = 27, p-value < 2.2e-16

2.2e-16 < 0.05

## [1] TRUE

### Respuesta. Los ingresos si son un factor importante para el consumo de libros.

<https://www.comunidadbaratz.com/blog/6-de-cada-10-personas-prefieren-leer-libros-en-papel/>

## pregunta 7

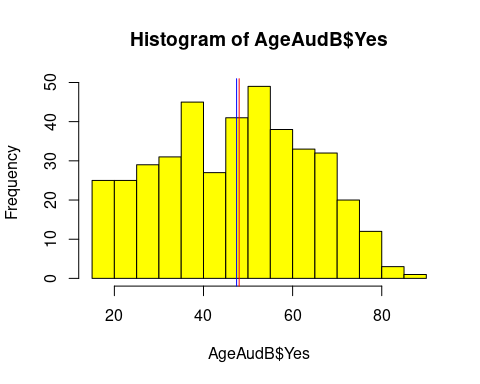
¿Los audio libros son frecuentados por jovenes o personas mayores?

Ho. Por ambos de una manera equitativa. H1. Por facilidad de uso y manejo asi que por personas mayores.

tabLec <- table(lectura$Age, lectura$Read.any.audiobooks.during.last.12months., dnn = c("Edad","Respuesta si escucho algun audio libro en el a;o"))  
 tabLec

## Respuesta si escucho algun audio libro en el a;o  
## Edad Don’t know No Yes  
## 16 2 0 55 11  
## 17 5 0 59 4  
## 18 5 0 50 4  
## 19 2 0 29 3  
## 20 2 0 29 3  
## 21 2 0 37 0  
## 22 3 0 33 9  
## 23 4 0 28 9  
## 24 4 0 25 5  
## 25 3 1 32 2  
## 26 6 0 24 1  
## 27 6 0 28 10  
## 28 7 0 16 5  
## 29 4 0 23 10  
## 30 8 0 39 3  
## 31 4 0 35 2  
## 32 8 0 35 9  
## 33 3 0 29 7  
## 34 11 0 25 4  
## 35 7 0 31 9  
## 36 3 0 24 8  
## 37 0 0 25 12  
## 38 1 0 28 7  
## 39 7 0 37 8  
## 40 7 0 43 10  
## 41 8 0 26 6  
## 42 3 0 38 7  
## 43 4 0 23 3  
## 44 8 0 33 3  
## 45 6 0 14 8  
## 46 8 0 19 6  
## 47 7 0 56 7  
## 48 5 0 24 11  
## 49 4 0 39 3  
## 50 13 0 51 14  
## 51 2 0 39 11  
## 52 6 0 36 11  
## 53 10 0 32 10  
## 54 10 1 39 3  
## 55 7 0 43 14  
## 56 7 0 35 5  
## 57 8 0 34 5  
## 58 8 0 39 6  
## 59 6 0 29 5  
## 60 8 0 47 17  
## 61 8 1 21 8  
## 62 6 0 35 7  
## 63 7 0 30 5  
## 64 4 0 29 6  
## 65 10 0 59 7  
## 66 6 0 35 10  
## 67 9 0 25 11  
## 68 9 0 22 4  
## 69 10 1 23 6  
## 70 8 0 29 1  
## 71 5 0 16 1  
## 72 6 0 14 6  
## 73 2 0 16 3  
## 74 2 0 11 4  
## 75 5 0 11 6  
## 76 3 0 11 3  
## 77 2 0 14 3  
## 78 4 0 20 3  
## 79 3 0 9 2  
## 80 4 0 16 1  
## 81 2 0 14 1  
## 82 6 0 10 0  
## 83 3 0 10 0  
## 84 5 0 5 2  
## 85 2 0 7 0  
## 86 0 0 3 1  
## 87 1 0 4 0  
## 88 2 0 7 0  
## 89 0 0 2 0  
## 90 1 0 1 0  
## 91 1 0 2 0  
## 92 2 0 0 0  
## 93 0 0 1 0

AgeAudB<- split(lectura$Age,lectura$Read.any.audiobooks.during.last.12months.)  
hist(AgeAudB$Yes, col = "yellow")  
  
mediaAudB <- mean(AgeAudB$Yes)  
abline(v=mediaAudB,col="blue")  
  
medianAudB <- median(AgeAudB$Yes)  
abline(v=medianAudB,col="red")



mediaAudB

## [1] 47.42822

medianAudB

## [1] 48

### Respuesta. Por personas mayores, segun nuestro grafico por personas entre los cuarenta y sesenta años mayoritariamente. Por tanto se rechaza la hipotesis nula, no hay equidad entre las edades y el hecho de escuchar audiolibros.

<https://computerhoy.com/noticias/industria/quien-lee-mejor-dicho-escucha-audiolibros-espana-405851>

## Pregunta 8.

De acuerdo a la pregunta anterior, habra normalidad en esos datos?

H0. Hay normalidad en los datos. H1. No hay normalidad (existe sesgo).

shapiro.test(AgeAudB$Yes)

##   
## Shapiro-Wilk normality test  
##   
## data: AgeAudB$Yes  
## W = 0.9788, p-value = 1.02e-05

1.02e-05 < 0.05

## [1] TRUE

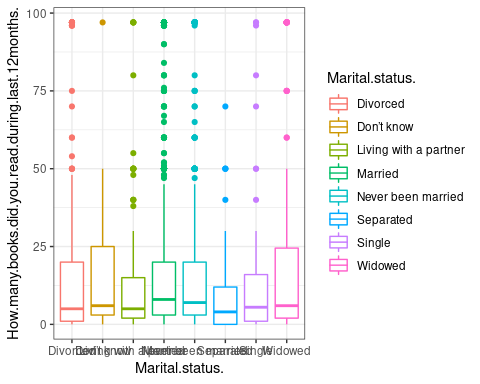
### Respuesta 2. Al buscar si hay normalidad en estos datos, nos damos cuenta que no la hay. Existe un sesgo hacia la derecha, señalando una mayoria de este lado, lo que quiere decir es que las personas que dijeron que si escuchan audiolibros son de edades despues de los 40 en su mayoria. Aunque hay una baja significativa despues de los 60s. Por lo tanto, se rechaza la hipotesis nula.

## pregunta 9.

Existe alguna diferencia entre los libros leidos de los casados con los solteros, y los divorciados?

H0. No hay diferencia H1. Si hay diferencia en la media de cada grupo

ggplot(data = lectura, aes(x= Marital.status. , y= How.many.books.did.you.read.during.last.12months., color= Marital.status.))+ geom\_boxplot()+theme\_bw()



anova <- aov(lectura$How.many.books.did.you.read.during.last.12months.~lectura$Marital.status.)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)  
## lectura$Marital.status. 7 5288 755.4 1.354 0.221  
## Residuals 2824 1575755 558.0

0.221 < 0.05

## [1] FALSE

### Respuesta.

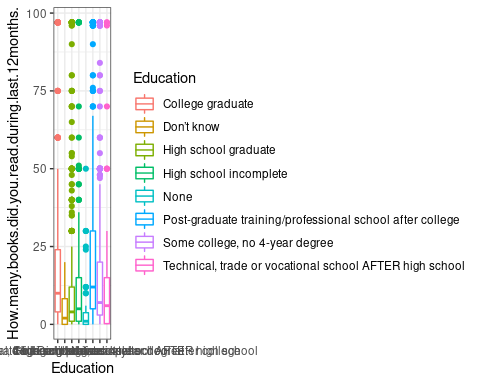
se rechaza la hipotesis nula por que si existe una diferencia entre los lectores y su estado civil. <https://www.cronista.com/informacion-gral/Los-hombres-y-mujeres-casados-leen-menos-que-los-solteros-20120424-0052.html>

## pregunta 10.

Influye el nivel educativo con respecto a si una persona lee o no?.

H0. No influye, hay equidad en los grupos. H1. Si influye, hay diferencia segun sea el nivel educativo.

ggplot(data = lectura, aes(x= Education , y= How.many.books.did.you.read.during.last.12months., color= Education))+ geom\_boxplot()+theme\_bw()



anova2 <- aov(lectura$How.many.books.did.you.read.during.last.12months.~lectura$Education)  
summary(anova2)

## Df Sum Sq Mean Sq F value Pr(>F)   
## lectura$Education 7 49718 7103 13.1 <2e-16 \*\*\*  
## Residuals 2824 1531324 542   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

2e-16 < 0

## [1] FALSE

### Respuesta.

Rechazamos la hipotesis nula y tomamos H1 por que si existe diferencia segun sea el nivel educativo. <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2020/EstSociodemo/MOLEC2019_04.pdf> <http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2218-36202018000300180#>:~:text=Por%20grupos%20de%20edad%20las,por%20conocer%20sobre%20alg%C3%BAn%20tema.

## Pregunta 11.

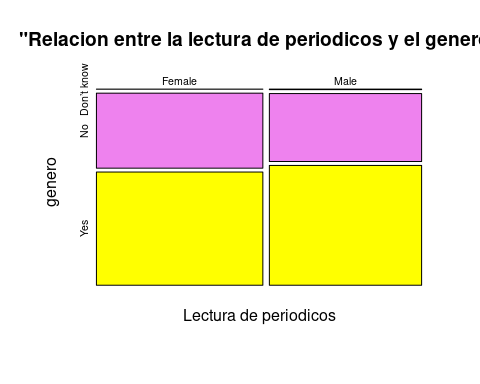
Los hombres leen mas el periodico que las mujeres?. Existe diferencia?

H0. No hay diferencia, la lectura de periodicos es independiente del genero H1. Si hay diferencia

tabGen <- table(lectura$Sex, lectura$Do.you.happen.to.read.any.daily.news.or.newspapers., dnn = c("Lectura de periodicos","genero"))  
 tabGen

## genero  
## Lectura de periodicos Don’t know No Yes  
## Female 1 589 889  
## Male 4 488 861

mosaicplot(tabGen, main = deparse("Relacion entre la lectura de periodicos y el genero"), color = c("yellow","violet"))



chisq.test(tabGen)

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

##   
## Pearson's Chi-squared test  
##   
## data: tabGen  
## X-squared = 6.1259, df = 2, p-value = 0.04675

0.04675 < 0.05

## [1] TRUE

### Respuesta.

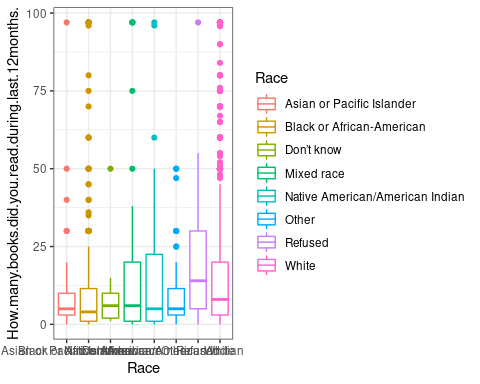
No rechazamos H0 ya que no hay diferencia, la lectura de periodicos es independiente del genero. <https://www.merca20.com/el-37-por-ciento-de-las-personas-que-tienen-entre-25-a-34-lee-mas-el-periodico/>

### Pregunta 12

¿Existe una diferencia en el habito de leer segun la etnia de la persona?

H0. No hay diferencia H1. Existe una diferencia

ggplot(data = lectura, aes(x= Race , y= How.many.books.did.you.read.during.last.12months., color= Race))+ geom\_boxplot()+theme\_bw()



anova3 <- aov(lectura$How.many.books.did.you.read.during.last.12months.~lectura$Race)  
summary(anova3)

## Df Sum Sq Mean Sq F value Pr(>F)   
## lectura$Race 7 24506 3501 6.352 2e-07 \*\*\*  
## Residuals 2824 1556536 551   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

2e-07 < 0

## [1] FALSE

### Respuesta.

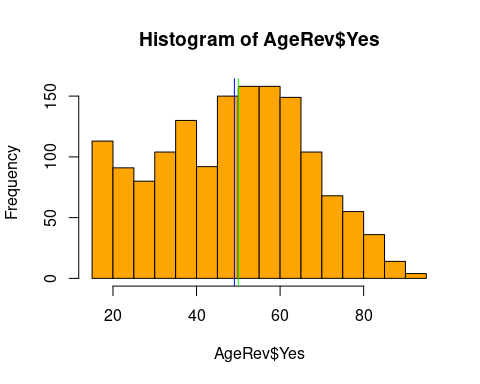
Se afirma h1 por que si existe una diferencia. Se rechaza h0. <https://revistasum.umanizales.edu.co/ojs/index.php/plumillaeducativa/article/view/1962/2625>

##pregunta 13 ¿la edad es un factor importante en la lectura de revistas o diarios?

tabMg <- table(lectura$Age, lectura$Do.you.happen.to.read.any.magazines.or.journals., dnn = c("Edad","Revistas o Diarios"))  
 tabMg

## Revistas o Diarios  
## Edad Don’t know No Yes  
## 16 1 36 31  
## 17 0 38 30  
## 18 0 34 25  
## 19 0 18 16  
## 20 0 23 11  
## 21 0 20 19  
## 22 0 24 21  
## 23 0 22 19  
## 24 0 23 11  
## 25 0 17 21  
## 26 0 23 8  
## 27 1 27 16  
## 28 0 19 9  
## 29 0 14 23  
## 30 0 26 24  
## 31 0 23 18  
## 32 0 28 24  
## 33 0 20 19  
## 34 0 22 18  
## 35 0 22 25  
## 36 0 13 22  
## 37 0 16 21  
## 38 0 14 22  
## 39 0 23 29  
## 40 0 24 36  
## 41 0 25 15  
## 42 0 29 19  
## 43 0 10 20  
## 44 0 22 22  
## 45 0 12 16  
## 46 0 16 17  
## 47 0 31 39  
## 48 0 15 25  
## 49 0 21 25  
## 50 0 34 44  
## 51 0 19 33  
## 52 0 19 34  
## 53 0 20 32  
## 54 0 24 29  
## 55 0 34 30  
## 56 0 22 25  
## 57 0 17 30  
## 58 0 19 34  
## 59 1 13 26  
## 60 0 29 43  
## 61 0 18 20  
## 62 0 23 25  
## 63 0 15 27  
## 64 0 15 24  
## 65 0 23 53  
## 66 1 25 25  
## 67 0 25 20  
## 68 0 14 21  
## 69 0 24 16  
## 70 0 16 22  
## 71 0 10 12  
## 72 0 11 15  
## 73 0 5 16  
## 74 0 7 10  
## 75 0 7 15  
## 76 0 9 8  
## 77 0 11 8  
## 78 0 13 14  
## 79 0 4 10  
## 80 0 6 15  
## 81 0 8 9  
## 82 0 6 10  
## 83 0 8 5  
## 84 0 7 5  
## 85 0 2 7  
## 86 0 2 2  
## 87 0 1 4  
## 88 0 4 5  
## 89 0 0 2  
## 90 0 1 1  
## 91 0 1 2  
## 92 0 1 1  
## 93 0 0 1

AgeRev<- split(lectura$Age,lectura$Do.you.happen.to.read.any.magazines.or.journals.)  
hist(AgeRev$Yes, col = "orange")  
mediaRev <- mean(AgeRev$Yes)  
abline(v=mediaRev,col="blue")  
  
medianRev <- median(AgeRev$Yes)  
abline(v=medianRev,col="green")



mediaRev

## [1] 49.05777

medianRev

## [1] 50

//según lo que encontre los jovenes leen más que los ancianos asi que no es un factor importante tambien. aqui te dejo el link de igual forma

<https://lapiedradesisifo.com/2016/04/27/los-millennials-leen-mas-que-los-mayores-de-30-anos/>

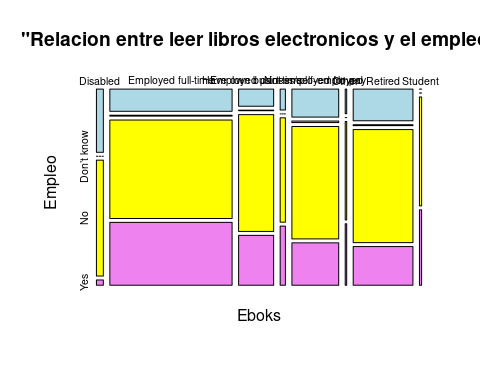
##pregunta 14 ¿Según el tipo de empleo de la persona aumenta el consumo de los libros electronicos?

Ho= Son variables independientes una de la otra (no hay relacion) H1 = no son variables independientes (si hay relacion)

tabE <- table(lectura$Employement, lectura$Read.any.e.books.during.last.12months., dnn = c("Eboks","Empleo"))  
 tabE

## Empleo  
## Eboks Don’t know No Yes  
## Disabled 24 0 44 2  
## Employed full-time 149 6 661 422  
## Employed part-time 33 1 225 96  
## Have own business/self-employed 6 0 30 17  
## Not employed for pay 72 4 289 109  
## Other 2 0 8 5  
## Retired 104 3 371 127  
## Student 0 0 13 9

mosaicplot(tabE, main = deparse("Relacion entre leer libros electronicos y el empleo"), color = c("lightblue","pink","yellow","violet"))

 ### Se puede apreciar que hay mayor consumo de libros electronicos en personas que tienen un empleo a tiempo completo, asi que suponemos esto se debe a las facilidades que adquisicion de este tipo o modalidad de libros.

chisq.test(tabE)

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

##   
## Pearson's Chi-squared test  
##   
## data: tabE  
## X-squared = 98.446, df = 21, p-value = 5.428e-12

5.428e-12 < 0.05

## [1] TRUE

### Respuesta. De acuerdo a lo que se puede ver en el grafico y la prueba de chi cuadrado, podermos decir que el consumo de libros electronicos si esta relacionado con el estatus de empleo que se tenga; por lo tanto se rechaza la hipotesis nula. Respondiendo a nuestra pregunta principal tenemos: segun lo encontrado lo prefieren por la comodidad. adjunto link

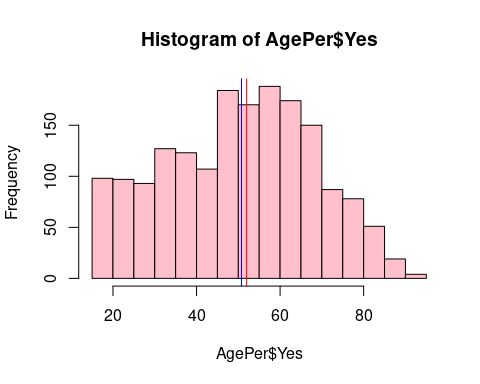
<http://www.revistaespacios.com/a18v39n43/a18v39n43p36.pdf>

##pregunta 15 ¿en base a la edad las personas prefieren leer solamente noticias en periodicos o diarios?

tabP <- table(lectura$Age, lectura$Do.you.happen.to.read.any.daily.news.or.newspapers., dnn = c("Edad","Revistas o Diarios"))  
 tabP

## Revistas o Diarios  
## Edad Don’t know No Yes  
## 16 0 41 27  
## 17 0 48 20  
## 18 0 35 24  
## 19 0 22 12  
## 20 0 19 15  
## 21 0 24 15  
## 22 0 25 20  
## 23 0 17 24  
## 24 0 20 14  
## 25 0 14 24  
## 26 0 16 15  
## 27 0 22 22  
## 28 1 13 14  
## 29 0 21 16  
## 30 0 24 26  
## 31 0 17 24  
## 32 0 30 22  
## 33 0 15 24  
## 34 1 12 27  
## 35 0 17 30  
## 36 0 12 23  
## 37 0 20 17  
## 38 0 12 24  
## 39 0 28 24  
## 40 0 25 35  
## 41 0 20 20  
## 42 0 18 30  
## 43 0 14 16  
## 44 0 19 25  
## 45 0 12 16  
## 46 0 7 26  
## 47 0 23 47  
## 48 0 15 25  
## 49 0 16 30  
## 50 1 21 56  
## 51 0 20 32  
## 52 0 18 35  
## 53 1 23 28  
## 54 0 20 33  
## 55 0 22 42  
## 56 0 15 32  
## 57 0 17 30  
## 58 0 15 38  
## 59 1 10 29  
## 60 0 13 59  
## 61 0 13 25  
## 62 0 15 33  
## 63 0 8 34  
## 64 0 12 27  
## 65 0 21 55  
## 66 0 10 41  
## 67 0 15 30  
## 68 0 10 25  
## 69 0 14 26  
## 70 0 10 28  
## 71 0 5 17  
## 72 0 7 19  
## 73 0 0 21  
## 74 0 4 13  
## 75 0 5 17  
## 76 0 4 13  
## 77 0 4 15  
## 78 0 7 20  
## 79 0 3 11  
## 80 0 2 19  
## 81 0 4 13  
## 82 0 6 10  
## 83 0 4 9  
## 84 0 1 11  
## 85 0 1 8  
## 86 0 0 4  
## 87 0 0 5  
## 88 0 2 7  
## 89 0 0 2  
## 90 0 1 1  
## 91 0 0 3  
## 92 0 1 1  
## 93 0 1 0

AgePer<- split(lectura$Age,lectura$Do.you.happen.to.read.any.daily.news.or.newspapers.)  
hist(AgePer$Yes, col = "pink")  
mediaPer <- mean(AgePer$Yes)  
abline(v=mediaPer,col="blue")  
  
medianPer <- median(AgePer$Yes)  
abline(v=medianPer,col="red")



mediaPer

## [1] 50.76057

medianPer

## [1] 52

### Respuesta. por comodiad tambien.

<https://www.larepublica.co/internet-economy/siete-de-cada-10-colombianos-aun-prefiere-leer-los-periodicos-impresos-2364096>

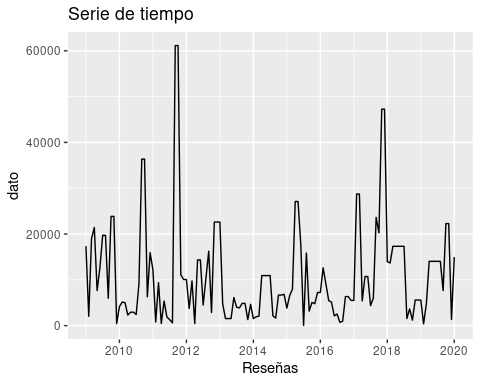
## Serie de tiempo de las reseñas de los libros más vendidos de amazon desde el 2009 al 2020

AmazonBooks <- read.csv("AmazonBooks - Sheet1.csv")

best2 = ts(AmazonBooks$Reviews, start = 2009, end = 2020, frequency = 12)  
best2

## Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec  
## 2009 17350 2052 18979 21424 7665 12643 19735 19699 5983 23848 23848 460  
## 2010 4149 5153 5013 2313 2925 2951 2426 9198 36348 36348 6310 15921  
## 2011 12159 798 9374 491 5360 1909 1296 615 61133 61133 11113 10070  
## 2012 10070 3729 9769 471 14344 14344 4505 10369 16244 2884 22614 22614  
## 2013 22614 4761 1542 1542 1542 6143 4022 3871 4866 4866 1329 4642  
## 2014 1541 1924 2094 10922 10922 10922 10922 2137 1651 6679 6679 6812  
## 2015 3837 6540 7955 27098 27098 17684 37 15845 3181 5062 4786 7235  
## 2016 7235 12619 9089 5470 5118 2134 2525 720 956 6346 6346 5505  
## 2017 5505 28729 28729 5413 10721 10721 4370 6042 23631 20262 47265 47265  
## 2018 13964 13677 17323 17323 17323 17323 17323 1555 3642 1215 5594 5594  
## 2019 5594 408 4799 14038 14038 14038 14038 14038 7660 22288 22288 1365  
## 2020 14982

autoplot(best2)+   
 labs(title = "Serie de tiempo",  
 x = "Reseñas",  
 y = "dato",  
colour = "#00a0dc")



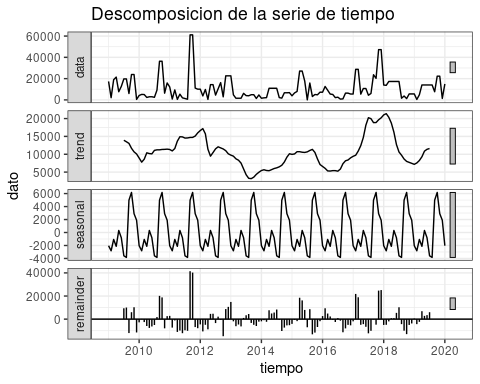
theme\_bw()

## List of 93  
## $ line :List of 6  
## ..$ colour : chr "black"  
## ..$ size : num 0.5  
## ..$ linetype : num 1  
## ..$ lineend : chr "butt"  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ rect :List of 5  
## ..$ fill : chr "white"  
## ..$ colour : chr "black"  
## ..$ size : num 0.5  
## ..$ linetype : num 1  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ text :List of 11  
## ..$ family : chr ""  
## ..$ face : chr "plain"  
## ..$ colour : chr "black"  
## ..$ size : num 11  
## ..$ hjust : num 0.5  
## ..$ vjust : num 0.5  
## ..$ angle : num 0  
## ..$ lineheight : num 0.9  
## ..$ margin : 'margin' num [1:4] 0points 0points 0points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ title : NULL  
## $ aspect.ratio : NULL  
## $ axis.title : NULL  
## $ axis.title.x :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 2.75points 0points 0points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.x.top :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 0  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 0points 2.75points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.x.bottom : NULL  
## $ axis.title.y :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 1  
## ..$ angle : num 90  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 2.75points 0points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.y.left : NULL  
## $ axis.title.y.right :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 0  
## ..$ angle : num -90  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 0points 0points 2.75points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : chr "grey30"  
## ..$ size : 'rel' num 0.8  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.x :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 2.2points 0points 0points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.x.top :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 0  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 0points 2.2points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.x.bottom : NULL  
## $ axis.text.y :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 1  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 2.2points 0points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.y.left : NULL  
## $ axis.text.y.right :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 0  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 0points 0points 2.2points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.ticks :List of 6  
## ..$ colour : chr "grey20"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ lineend : NULL  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ axis.ticks.x : NULL  
## $ axis.ticks.x.top : NULL  
## $ axis.ticks.x.bottom : NULL  
## $ axis.ticks.y : NULL  
## $ axis.ticks.y.left : NULL  
## $ axis.ticks.y.right : NULL  
## $ axis.ticks.length : 'simpleUnit' num 2.75points  
## ..- attr(\*, "unit")= int 8  
## $ axis.ticks.length.x : NULL  
## $ axis.ticks.length.x.top : NULL  
## $ axis.ticks.length.x.bottom: NULL  
## $ axis.ticks.length.y : NULL  
## $ axis.ticks.length.y.left : NULL  
## $ axis.ticks.length.y.right : NULL  
## $ axis.line : list()  
## ..- attr(\*, "class")= chr [1:2] "element\_blank" "element"  
## $ axis.line.x : NULL  
## $ axis.line.x.top : NULL  
## $ axis.line.x.bottom : NULL  
## $ axis.line.y : NULL  
## $ axis.line.y.left : NULL  
## $ axis.line.y.right : NULL  
## $ legend.background :List of 5  
## ..$ fill : NULL  
## ..$ colour : logi NA  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ legend.margin : 'margin' num [1:4] 5.5points 5.5points 5.5points 5.5points  
## ..- attr(\*, "unit")= int 8  
## $ legend.spacing : 'simpleUnit' num 11points  
## ..- attr(\*, "unit")= int 8  
## $ legend.spacing.x : NULL  
## $ legend.spacing.y : NULL  
## $ legend.key :List of 5  
## ..$ fill : chr "white"  
## ..$ colour : logi NA  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ legend.key.size : 'simpleUnit' num 1.2lines  
## ..- attr(\*, "unit")= int 3  
## $ legend.key.height : NULL  
## $ legend.key.width : NULL  
## $ legend.text :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : 'rel' num 0.8  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ legend.text.align : NULL  
## $ legend.title :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 0  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ legend.title.align : NULL  
## $ legend.position : chr "right"  
## $ legend.direction : NULL  
## $ legend.justification : chr "center"  
## $ legend.box : NULL  
## $ legend.box.just : NULL  
## $ legend.box.margin : 'margin' num [1:4] 0cm 0cm 0cm 0cm  
## ..- attr(\*, "unit")= int 1  
## $ legend.box.background : list()  
## ..- attr(\*, "class")= chr [1:2] "element\_blank" "element"  
## $ legend.box.spacing : 'simpleUnit' num 11points  
## ..- attr(\*, "unit")= int 8  
## $ panel.background :List of 5  
## ..$ fill : chr "white"  
## ..$ colour : logi NA  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ panel.border :List of 5  
## ..$ fill : logi NA  
## ..$ colour : chr "grey20"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ panel.spacing : 'simpleUnit' num 5.5points  
## ..- attr(\*, "unit")= int 8  
## $ panel.spacing.x : NULL  
## $ panel.spacing.y : NULL  
## $ panel.grid :List of 6  
## ..$ colour : chr "grey92"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ lineend : NULL  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ panel.grid.major : NULL  
## $ panel.grid.minor :List of 6  
## ..$ colour : NULL  
## ..$ size : 'rel' num 0.5  
## ..$ linetype : NULL  
## ..$ lineend : NULL  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ panel.grid.major.x : NULL  
## $ panel.grid.major.y : NULL  
## $ panel.grid.minor.x : NULL  
## $ panel.grid.minor.y : NULL  
## $ panel.ontop : logi FALSE  
## $ plot.background :List of 5  
## ..$ fill : NULL  
## ..$ colour : chr "white"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ plot.title :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : 'rel' num 1.2  
## ..$ hjust : num 0  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 0points 5.5points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.title.position : chr "panel"  
## $ plot.subtitle :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 0  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 0points 0points 5.5points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.caption :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : 'rel' num 0.8  
## ..$ hjust : num 1  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 5.5points 0points 0points 0points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.caption.position : chr "panel"  
## $ plot.tag :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : 'rel' num 1.2  
## ..$ hjust : num 0.5  
## ..$ vjust : num 0.5  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.tag.position : chr "topleft"  
## $ plot.margin : 'margin' num [1:4] 5.5points 5.5points 5.5points 5.5points  
## ..- attr(\*, "unit")= int 8  
## $ strip.background :List of 5  
## ..$ fill : chr "grey85"  
## ..$ colour : chr "grey20"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ strip.background.x : NULL  
## $ strip.background.y : NULL  
## $ strip.placement : chr "inside"  
## $ strip.text :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : chr "grey10"  
## ..$ size : 'rel' num 0.8  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : 'margin' num [1:4] 4.4points 4.4points 4.4points 4.4points  
## .. ..- attr(\*, "unit")= int 8  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ strip.text.x : NULL  
## $ strip.text.y :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : num -90  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ strip.switch.pad.grid : 'simpleUnit' num 2.75points  
## ..- attr(\*, "unit")= int 8  
## $ strip.switch.pad.wrap : 'simpleUnit' num 2.75points  
## ..- attr(\*, "unit")= int 8  
## $ strip.text.y.left :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : num 90  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## - attr(\*, "class")= chr [1:2] "theme" "gg"  
## - attr(\*, "complete")= logi TRUE  
## - attr(\*, "validate")= logi TRUE

### Segun lo que podemos ver en la grafica hubo un periodo de mayor intesidad con respecto a las reseñas en el año 2012, y otro en el 2018, sin embargo los otros años se mantuvo de alguna forma estable y algo bajo el numero de reseñas escritas en Amazon.

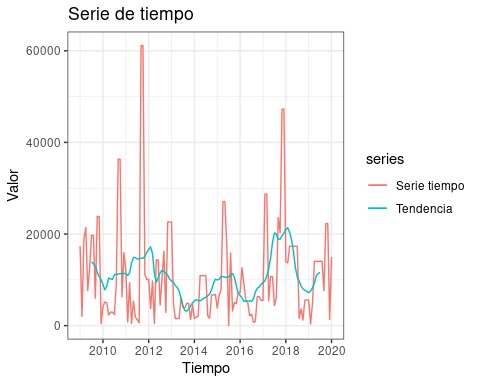
## Descomposicion de serie de tiempo

D <- decompose(best2, type = 'additive')  
  
autoplot(D)+  
 labs(title= "Descomposicion de la serie de tiempo",   
 x= "tiempo",  
 y = "dato",  
 colour = "Gears")+  
 theme\_bw()

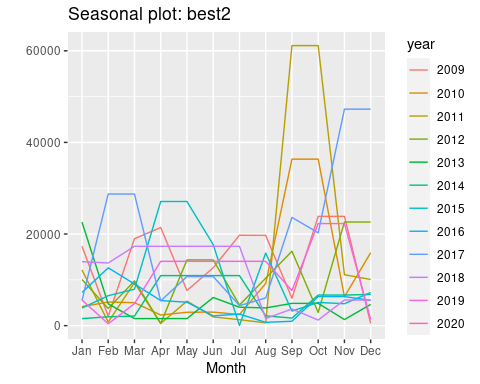
 ### Tendencia

autoplot(best2, series="Serie tiempo") +   
 autolayer(trendcycle(D), series="Tendencia") +  
 labs(title = "Serie de tiempo",   
 x = "Tiempo",  
 y = "Valor"  
 ) +   
 theme\_bw()

## Warning: Removed 12 row(s) containing missing values (geom\_path).

 ### Para ver estacionalidad

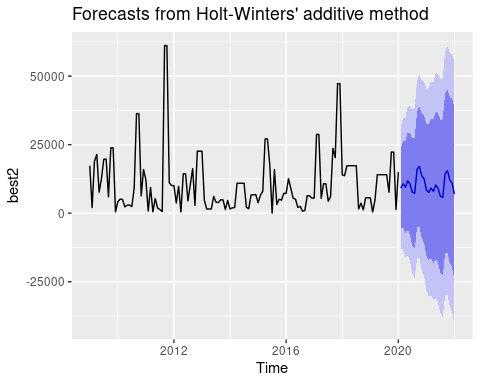
ggseasonplot(best2)

 ## Pronostico de la Serie de Tiempo

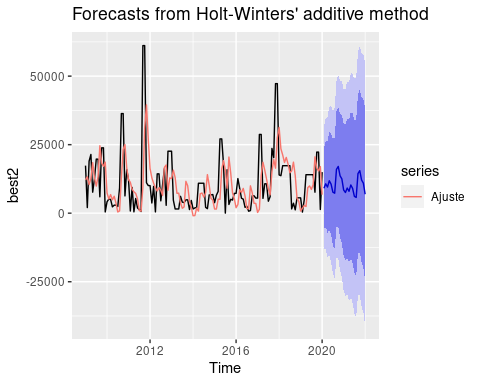
### Método holt winters

# elaborando el pronostico  
most <- hw(best2, h=24, seasonal = 'additive')

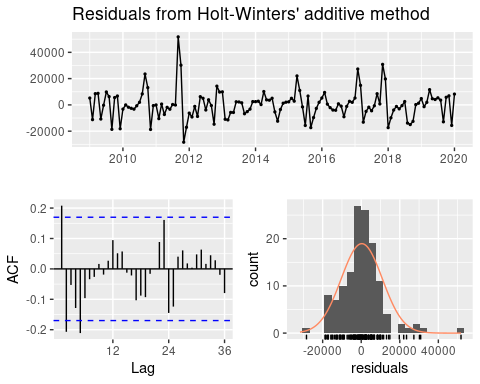
# graficando el pronóstico  
autoplot(most)



# verificando el ajuste del método  
autoplot(most)+autolayer(fitted(most), series="Ajuste")



# verificando los residuales  
checkresiduals(most)



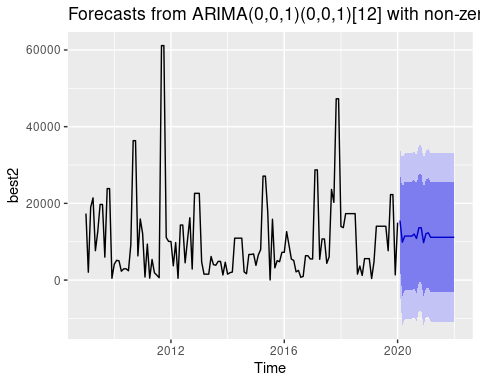
##   
## Ljung-Box test  
##   
## data: Residuals from Holt-Winters' additive method  
## Q\* = 37.972, df = 8, p-value = 7.618e-06  
##   
## Model df: 16. Total lags used: 24

### Método Arima

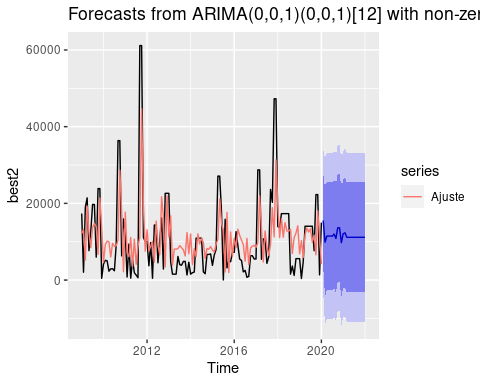
# elaborando el modelo ARIMA  
modeloAr <- auto.arima(best2)

# elaborando el pronostico  
most2 <- forecast(modeloAr, h=24)

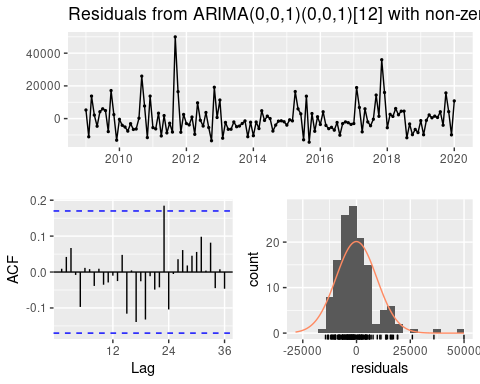
# graficando el pronóstico  
autoplot(most2)



# verificando el ajuste del método  
autoplot(most2)+autolayer(fitted(most2), series="Ajuste")



# verificando los residuales  
checkresiduals(most2)



##   
## Ljung-Box test  
##   
## data: Residuals from ARIMA(0,0,1)(0,0,1)[12] with non-zero mean  
## Q\* = 19.171, df = 21, p-value = 0.5742  
##   
## Model df: 3. Total lags used: 24

##adjuntamos links sobre el porque amazon hace reseñas sobre sus productos. <https://sellercentral.amazon.com.mx/gp/help/external/201972160?language=es_MX&ref=efph_201972160_cont_201972140#>:~:text=Amazon%20muestra%20las%20rese%C3%B1as%20de,hora%20de%20comprar%20un%20producto.

<https://sellercentral.amazon.es/gp/help/external/201972160?language=es_ES&ref=efph_201972160_cont_201972140>