Examen Final Estadistica II

Arnold David Bernabe Jeronimo Mijangos

3/6/2021

## Librerias

require(stats)  
library(dplyr)

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

##   
## 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(readr)  
library(normtest)  
library(nortest)  
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 stringr 1.4.0  
## v tidyr 1.1.3 v forcats 0.5.1  
## v purrr 0.3.4

## 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 '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()

## Datasets a Utilizar

Auts <- read\_csv("C:/Users/David/Documents/Proyectos/ExamFinalEst/AutosSegundaMano.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## .default = col\_character(),  
## price = col\_double(),  
## price\_financed = col\_double(),  
## year = col\_double(),  
## kms = col\_double(),  
## power = col\_double(),  
## doors = col\_double(),  
## photos = col\_double(),  
## is\_professional = col\_logical(),  
## publish\_date = col\_datetime(format = ""),  
## insert\_date = col\_datetime(format = "")  
## )  
## i Use `spec()` for the full column specifications.

Netflix <- read\_csv("C:/Users/David/Documents/Proyectos/ExamFinalEst/netflix.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## .default = col\_character(),  
## `Hidden Gem Score` = col\_double(),  
## `IMDb Score` = col\_double(),  
## `Rotten Tomatoes Score` = col\_double(),  
## `Metacritic Score` = col\_double(),  
## `Awards Received` = col\_double(),  
## `Awards Nominated For` = col\_double(),  
## `Netflix Release Date` = col\_date(format = ""),  
## `IMDb Votes` = col\_double()  
## )  
## i Use `spec()` for the full column specifications.

Co2Est <- read\_csv("C:/Users/David/Documents/Proyectos/ExamFinalEst/archive.csv")

##   
## -- Column specification --------------------------------------------------------  
## cols(  
## Year = col\_double(),  
## Month = col\_double(),  
## `Decimal Date` = col\_double(),  
## `Carbon Dioxide (ppm)` = col\_double(),  
## `Seasonally Adjusted CO2 (ppm)` = col\_double(),  
## `Carbon Dioxide Fit (ppm)` = col\_double(),  
## `Seasonally Adjusted CO2 Fit (ppm)` = col\_double()  
## )

## Analisis exploratorio de los datos

## Dataset AutosSegundaManos.csv

Para el dataset de Autos de segunda mano en venta tenemos los siguientes datos:

str(Auts)

## spec\_tbl\_df [50,000 x 21] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ url : chr [1:50000] "e158ae0ca53119ca199c28c36b5c2fcd" "ff267ebb7e700246f47f84f3db660b4b" "de4b02db28ea7786c622b969be10c7c7" "0449972a4d07594acf92e9a7dd28b39c" ...  
## $ company : chr [1:50000] "9881bcdd5a0ad4733037b3fb25e69c3a" "9881bcdd5a0ad4733037b3fb25e69c3a" "9881bcdd5a0ad4733037b3fb25e69c3a" "9881bcdd5a0ad4733037b3fb25e69c3a" ...  
## $ make : chr [1:50000] "SEAT" "CITROEN" "FORD" "VOLKSWAGEN" ...  
## $ model : chr [1:50000] "Toledo" "C1" "Transit Connect" "Caravelle" ...  
## $ version : chr [1:50000] "SEAT Toledo 4p." "CITROEN C1 PureTech 60KW 82CV Feel 5p." "FORD Transit Connect Van 1.5 TDCi 100cv Ambiente 200 L1" "VOLKSWAGEN Caravelle Largo 2.0 TDI 140 Comfortlin Edition BMT" ...  
## $ price : num [1:50000] 950 6200 7851 19426 22850 ...  
## $ price\_financed : num [1:50000] NA NA 7024 NA 22800 ...  
## $ fuel : chr [1:50000] "Diésel" "Gasolina" "Diésel" "Diésel" ...  
## $ year : num [1:50000] 2000 2017 2016 2014 2017 ...  
## $ kms : num [1:50000] 227000 50071 103000 120000 107000 ...  
## $ power : num [1:50000] NA 82 100 140 130 130 150 150 110 80 ...  
## $ doors : num [1:50000] 4 5 4 4 2 5 5 5 5 5 ...  
## $ shift : chr [1:50000] "Manual" "Manual" "Manual" "Manual" ...  
## $ color : chr [1:50000] "Verde" "Blanco" "Blanco" "Blanco" ...  
## $ photos : num [1:50000] 5 6 10 9 4 32 47 15 6 6 ...  
## $ is\_professional: logi [1:50000] FALSE TRUE TRUE TRUE TRUE TRUE ...  
## $ dealer : chr [1:50000] "0f4bb8455d27349b8273109b66a847f3" "Autos Raymara" "Auto 96" "Inniauto" ...  
## $ province : chr [1:50000] "Navarra" "Tenerife" "Barcelona" "Navarra" ...  
## $ country : chr [1:50000] "Spain" "Spain" "Spain" "Spain" ...  
## $ publish\_date : POSIXct[1:50000], format: "2020-12-18 10:47:13" "2021-01-02 11:25:40" ...  
## $ insert\_date : POSIXct[1:50000], format: "2021-01-15" "2021-01-15" ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. url = col\_character(),  
## .. company = col\_character(),  
## .. make = col\_character(),  
## .. model = col\_character(),  
## .. version = col\_character(),  
## .. price = col\_double(),  
## .. price\_financed = col\_double(),  
## .. fuel = col\_character(),  
## .. year = col\_double(),  
## .. kms = col\_double(),  
## .. power = col\_double(),  
## .. doors = col\_double(),  
## .. shift = col\_character(),  
## .. color = col\_character(),  
## .. photos = col\_double(),  
## .. is\_professional = col\_logical(),  
## .. dealer = col\_character(),  
## .. province = col\_character(),  
## .. country = col\_character(),  
## .. publish\_date = col\_datetime(format = ""),  
## .. insert\_date = col\_datetime(format = "")  
## .. )

head(Auts)

## # A tibble: 6 x 21  
## url company make model version price price\_financed fuel year kms  
## <chr> <chr> <chr> <chr> <chr> <dbl> <dbl> <chr> <dbl> <dbl>  
## 1 e158ae~ 9881bcd~ SEAT Tole~ SEAT Tol~ 950 NA Diés~ 2000 227000  
## 2 ff267e~ 9881bcd~ CITR~ C1 CITROEN ~ 6200 NA Gaso~ 2017 50071  
## 3 de4b02~ 9881bcd~ FORD Tran~ FORD Tra~ 7851 7024 Diés~ 2016 103000  
## 4 044997~ 9881bcd~ VOLK~ Cara~ VOLKSWAG~ 19426 NA Diés~ 2014 120000  
## 5 12c4fa~ 9881bcd~ FORD Tran~ FORD Tra~ 22850 22800 Diés~ 2017 107000  
## 6 f15e9a~ 9881bcd~ PEUG~ 3008 PEUGEOT ~ 11490 10490 Gaso~ 2016 78665  
## # ... with 11 more variables: power <dbl>, doors <dbl>, shift <chr>,  
## # color <chr>, photos <dbl>, is\_professional <lgl>, dealer <chr>,  
## # province <chr>, country <chr>, publish\_date <dttm>, insert\_date <dttm>

summary(Auts)

## url company make model   
## Length:50000 Length:50000 Length:50000 Length:50000   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## version price price\_financed fuel   
## Length:50000 Min. : 300 Min. : 700 Length:50000   
## Class :character 1st Qu.: 5500 1st Qu.: 10790 Class :character   
## Mode :character Median : 11750 Median : 15500 Mode :character   
## Mean : 15166 Mean : 18480   
## 3rd Qu.: 19800 3rd Qu.: 21995   
## Max. :549900 Max. :428999   
## NA's :26437   
## year kms power doors   
## Min. :1970 Min. : 0 Min. : 29.0 Min. :2.000   
## 1st Qu.:2007 1st Qu.: 38300 1st Qu.:109.0 1st Qu.:4.000   
## Median :2015 Median : 99999 Median :130.0 Median :5.000   
## Mean :2012 Mean : 116359 Mean :146.4 Mean :4.505   
## 3rd Qu.:2018 3rd Qu.: 174900 3rd Qu.:163.0 3rd Qu.:5.000   
## Max. :2021 Max. :5000000 Max. :800.0 Max. :5.000   
## NA's :2 NA's :8528   
## shift color photos is\_professional  
## Length:50000 Length:50000 Min. : 0.0 Mode :logical   
## Class :character Class :character 1st Qu.: 7.0 FALSE:17451   
## Mode :character Mode :character Median :13.0 TRUE :32549   
## Mean :16.2   
## 3rd Qu.:23.0   
## Max. :54.0   
##   
## dealer province country   
## Length:50000 Length:50000 Length:50000   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## publish\_date insert\_date   
## Min. :2017-05-09 14:56:20 Min. :2021-01-15   
## 1st Qu.:2020-12-10 21:08:38 1st Qu.:2021-01-15   
## Median :2020-12-30 16:17:31 Median :2021-01-15   
## Mean :2020-12-21 17:22:53 Mean :2021-01-15   
## 3rd Qu.:2021-01-09 21:27:41 3rd Qu.:2021-01-15   
## Max. :2021-01-15 12:40:51 Max. :2021-01-15   
##

En el cual podemos observar que de parte de las variables cualitativas tenemos: Url, Company, make, model, version, fuel, shift, color, dealer, province, country.

En el apartado de las variables cuantitativas tenemos: price, price\_financed, year, kms, power, doors, photos.

y tambien podemos observar que contamos con variables de tipo fecha las cuales son: publish\_date e instert\_date.

## Conclusion:

podemos observar que se cuenta con un dataset variado, el cual presenta tanto variables cualitativas como cuantitativas y de fecha, por lo tanto aun teniendo variables inutilizables para este analisis como lo son Url y Company podemos analizar datos relevantes.

## Dataset netflix.csv

Para el dataset de Catalogo de netflix y su metascore tenemos los siguientes datos:

str(Netflix)

## spec\_tbl\_df [15,480 x 29] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ Title : chr [1:15480] "Lets Fight Ghost" "HOW TO BUILD A GIRL" "Centigrade" "ANNE+" ...  
## $ Genre : chr [1:15480] "Crime, Drama, Fantasy, Horror, Romance" "Comedy" "Drama, Thriller" "Drama" ...  
## $ Tags : chr [1:15480] "Comedy Programmes,Romantic TV Comedies,Horror Programmes,Thai TV Programmes" "Dramas,Comedies,Films Based on Books,British" "Thrillers" "TV Dramas,Romantic TV Dramas,Dutch TV Shows" ...  
## $ Languages : chr [1:15480] "Swedish, Spanish" "English" "English" "Turkish" ...  
## $ Series or Movie : chr [1:15480] "Series" "Movie" "Movie" "Series" ...  
## $ Hidden Gem Score : num [1:15480] 4.3 7 6.4 7.7 8.1 8.6 8.7 6.9 8.3 5.3 ...  
## $ Country Availability : chr [1:15480] "Thailand" "Canada" "Canada" "Belgium,Netherlands" ...  
## $ Runtime : chr [1:15480] "< 30 minutes" "1-2 hour" "1-2 hour" "< 30 minutes" ...  
## $ Director : chr [1:15480] "Tomas Alfredson" "Coky Giedroyc" "Brendan Walsh" NA ...  
## $ Writer : chr [1:15480] "John Ajvide Lindqvist" "Caitlin Moran" "Brendan Walsh, Daley Nixon" NA ...  
## $ Actors : chr [1:15480] "Kåre Hedebrant, Per Ragnar, Lina Leandersson, Henrik Dahl" "Paddy Considine, Cleo, Beanie Feldstein, Dónal Finn" "Genesis Rodriguez, Vincent Piazza" "Vahide Perçin, Gonca Vuslateri, Cansu Dere, Beren Gokyildiz" ...  
## $ View Rating : chr [1:15480] "R" "R" "Unrated" NA ...  
## $ IMDb Score : num [1:15480] 7.9 5.8 4.3 6.5 6.3 7.4 7.5 3.9 6.7 6.6 ...  
## $ Rotten Tomatoes Score: num [1:15480] 98 79 NA NA NA NA NA NA NA NA ...  
## $ Metacritic Score : num [1:15480] 82 69 46 NA NA NA NA NA NA NA ...  
## $ Awards Received : num [1:15480] 74 1 NA 1 NA NA 2 NA 2 NA ...  
## $ Awards Nominated For : num [1:15480] 57 NA NA NA 4 NA 4 NA 1 NA ...  
## $ Boxoffice : chr [1:15480] "$2,122,065" "$70,632" "$16,263" NA ...  
## $ Release Date : chr [1:15480] "12 Dec 2008" "08 May 2020" "28 Aug 2020" "01 Oct 2016" ...  
## $ Netflix Release Date : Date[1:15480], format: "2021-03-04" "2021-03-04" ...  
## $ Production House : chr [1:15480] "Canal+, Sandrew Metronome" "Film 4, Monumental Pictures, Lionsgate" NA NA ...  
## $ Netflix Link : chr [1:15480] "https://www.netflix.com/watch/81415947" "https://www.netflix.com/watch/81041267" "https://www.netflix.com/watch/81305978" "https://www.netflix.com/watch/81336456" ...  
## $ IMDb Link : chr [1:15480] "https://www.imdb.com/title/tt1139797" "https://www.imdb.com/title/tt4193072" "https://www.imdb.com/title/tt8945942" "https://www.imdb.com/title/tt6132758" ...  
## $ Summary : chr [1:15480] "A med student with a supernatural gift tries to cash in on his abilities by facing off against ghosts, till a w"| \_\_truncated\_\_ "When nerdy Johanna moves to London, things get out of hand when she reinvents herself as a bad-mouthed music cr"| \_\_truncated\_\_ "Trapped in a frozen car during a blizzard, a pregnant woman and her husband fight to survive while the temperat"| \_\_truncated\_\_ "Upon moving into a new place, a 20-something runs into a former flame that triggers memories of past relationsh"| \_\_truncated\_\_ ...  
## $ IMDb Votes : num [1:15480] 205926 2838 1720 1147 63 ...  
## $ Image : chr [1:15480] "https://occ-0-4708-64.1.nflxso.net/dnm/api/v6/evlCitJPPCVCry0BZlEFb5-QjKc/AAAABcmgLCxN8dNahdY2kgd1hhcL2a6XrE92x"| \_\_truncated\_\_ "https://occ-0-1081-999.1.nflxso.net/dnm/api/v6/evlCitJPPCVCry0BZlEFb5-QjKc/AAAABe\_fxMSBM1E-sSoszr12SmkI-498sqBW"| \_\_truncated\_\_ "https://occ-0-1081-999.1.nflxso.net/dnm/api/v6/evlCitJPPCVCry0BZlEFb5-QjKc/AAAABW-fG-2\_s3pGsBdbw4nLCWENcRKL2Ngj"| \_\_truncated\_\_ "https://occ-0-1489-1490.1.nflxso.net/dnm/api/v6/evlCitJPPCVCry0BZlEFb5-QjKc/AAAABeq3p0f50KPNQTYmozdPUenqXI3bh6H"| \_\_truncated\_\_ ...  
## $ Poster : chr [1:15480] "https://m.media-amazon.com/images/M/MV5BOWM4NTY2NTMtZDZlZS00NTgyLWEzZDMtODE3ZGI1MzI3ZmU5XkEyXkFqcGdeQXVyNzI1NzM"| \_\_truncated\_\_ "https://m.media-amazon.com/images/M/MV5BZGUyN2ZlMjYtZTk2Yy00MWZiLWIyMDktMzFlMmEzOWVlMGNiXkEyXkFqcGdeQXVyMTE1MzI"| \_\_truncated\_\_ "https://m.media-amazon.com/images/M/MV5BODM2MDQ5MjktYTM1ZS00Y2M4LTg0MDAtZjFjZDM1MDQxZGRmXkEyXkFqcGdeQXVyNDA1NDA"| \_\_truncated\_\_ "https://m.media-amazon.com/images/M/MV5BNWRkMzdiYjgtOTA0Yi00NjZiLWFjZjMtYThlMTE5MWEwYWU4XkEyXkFqcGdeQXVyODY1MDk"| \_\_truncated\_\_ ...  
## $ TMDb Trailer : chr [1:15480] NA "https://www.youtube.com/watch?v=eIbcxPy4okQ" "https://www.youtube.com/watch?v=0RvV7TNUlkQ" NA ...  
## $ Trailer Site : chr [1:15480] NA "YouTube" "YouTube" NA ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. Title = col\_character(),  
## .. Genre = col\_character(),  
## .. Tags = col\_character(),  
## .. Languages = col\_character(),  
## .. `Series or Movie` = col\_character(),  
## .. `Hidden Gem Score` = col\_double(),  
## .. `Country Availability` = col\_character(),  
## .. Runtime = col\_character(),  
## .. Director = col\_character(),  
## .. Writer = col\_character(),  
## .. Actors = col\_character(),  
## .. `View Rating` = col\_character(),  
## .. `IMDb Score` = col\_double(),  
## .. `Rotten Tomatoes Score` = col\_double(),  
## .. `Metacritic Score` = col\_double(),  
## .. `Awards Received` = col\_double(),  
## .. `Awards Nominated For` = col\_double(),  
## .. Boxoffice = col\_character(),  
## .. `Release Date` = col\_character(),  
## .. `Netflix Release Date` = col\_date(format = ""),  
## .. `Production House` = col\_character(),  
## .. `Netflix Link` = col\_character(),  
## .. `IMDb Link` = col\_character(),  
## .. Summary = col\_character(),  
## .. `IMDb Votes` = col\_double(),  
## .. Image = col\_character(),  
## .. Poster = col\_character(),  
## .. `TMDb Trailer` = col\_character(),  
## .. `Trailer Site` = col\_character()  
## .. )

head(Netflix)

## # A tibble: 6 x 29  
## Title Genre Tags Languages `Series or Movi~ `Hidden Gem Sco~ `Country Availa~  
## <chr> <chr> <chr> <chr> <chr> <dbl> <chr>   
## 1 Lets~ Crim~ Come~ Swedish,~ Series 4.3 Thailand   
## 2 HOW ~ Come~ Dram~ English Movie 7 Canada   
## 3 Cent~ Dram~ Thri~ English Movie 6.4 Canada   
## 4 ANNE+ Drama TV D~ Turkish Series 7.7 Belgium,Netherl~  
## 5 Moxie Anim~ Soci~ English Movie 8.1 Lithuania,Polan~  
## 6 The ~ Come~ Roma~ Thai Movie 8.6 Thailand   
## # ... with 22 more variables: Runtime <chr>, Director <chr>, Writer <chr>,  
## # Actors <chr>, View Rating <chr>, IMDb Score <dbl>,  
## # Rotten Tomatoes Score <dbl>, Metacritic Score <dbl>, Awards Received <dbl>,  
## # Awards Nominated For <dbl>, Boxoffice <chr>, Release Date <chr>,  
## # Netflix Release Date <date>, Production House <chr>, Netflix Link <chr>,  
## # IMDb Link <chr>, Summary <chr>, IMDb Votes <dbl>, Image <chr>,  
## # Poster <chr>, TMDb Trailer <chr>, Trailer Site <chr>

summary(Netflix)

## Title Genre Tags Languages   
## Length:15480 Length:15480 Length:15480 Length:15480   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Series or Movie Hidden Gem Score Country Availability Runtime   
## Length:15480 Min. :0.600 Length:15480 Length:15480   
## Class :character 1st Qu.:3.800 Class :character Class :character   
## Mode :character Median :6.800 Mode :character Mode :character   
## Mean :5.938   
## 3rd Qu.:7.900   
## Max. :9.800   
## NA's :2101   
## Director Writer Actors View Rating   
## Length:15480 Length:15480 Length:15480 Length:15480   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## IMDb Score Rotten Tomatoes Score Metacritic Score Awards Received   
## Min. :1.000 Min. : 0.00 Min. : 5.00 Min. : 1.000   
## 1st Qu.:5.800 1st Qu.: 38.00 1st Qu.: 44.00 1st Qu.: 1.000   
## Median :6.600 Median : 64.00 Median : 57.00 Median : 3.000   
## Mean :6.496 Mean : 59.52 Mean : 56.81 Mean : 8.764   
## 3rd Qu.:7.300 3rd Qu.: 83.00 3rd Qu.: 70.00 3rd Qu.: 8.000   
## Max. :9.700 Max. :100.00 Max. :100.00 Max. :300.000   
## NA's :2099 NA's :9098 NA's :11144 NA's :9405   
## Awards Nominated For Boxoffice Release Date   
## Min. : 1.00 Length:15480 Length:15480   
## 1st Qu.: 2.00 Class :character Class :character   
## Median : 5.00 Mode :character Mode :character   
## Mean : 13.98   
## 3rd Qu.: 12.00   
## Max. :386.00   
## NA's :7819   
## Netflix Release Date Production House Netflix Link IMDb Link   
## Min. :2015-04-14 Length:15480 Length:15480 Length:15480   
## 1st Qu.:2016-08-09 Class :character Class :character Class :character   
## Median :2018-10-05 Mode :character Mode :character Mode :character   
## Mean :2018-05-18   
## 3rd Qu.:2020-03-18   
## Max. :2021-03-04   
##   
## Summary IMDb Votes Image Poster   
## Length:15480 Min. : 5.0 Length:15480 Length:15480   
## Class :character 1st Qu.: 403.5 Class :character Class :character   
## Mode :character Median : 2322.0 Mode :character Mode :character   
## Mean : 42728.4   
## 3rd Qu.: 20890.5   
## Max. :2354197.0   
## NA's :2101   
## TMDb Trailer Trailer Site   
## Length:15480 Length:15480   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##   
##

En el cual podemos observar que de parte de las variables cualitativas tenemos: Title, Genre, Tags, languages, Series or Movie, Country Availability, Runtime, Director, Writer, Actors, View Rating, Boxoffice, release date, production House, netflix link, imdb link, sumary, Image, Poster, imdb trailer, trailer Site.

En el apartado de las variables cuantitativas tenemos: Hiden Gem Score, IMDb Score, Rotten tomatoes Score, Awards Reicivedm Awards nominated for, Imdb votes.

y tambien podemos observar que contamos con una variable de tipo fecha la cual es Netflix Release Date, la cual para este analisis no sera de mucha relevancia ya que la fecha es de cuando se tomaron los datos por lo cual no tiene mayor repercusion en el set de datos.

## Dataset Co2Est

se realizo la inclusion de este dataset para desarrollar de la mejor manera las series de tiempo

str(Co2Est)

## spec\_tbl\_df [720 x 7] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ Year : num [1:720] 1958 1958 1958 1958 1958 ...  
## $ Month : num [1:720] 1 2 3 4 5 6 7 8 9 10 ...  
## $ Decimal Date : num [1:720] 1958 1958 1958 1958 1958 ...  
## $ Carbon Dioxide (ppm) : num [1:720] NA NA 316 317 318 ...  
## $ Seasonally Adjusted CO2 (ppm) : num [1:720] NA NA 314 315 315 ...  
## $ Carbon Dioxide Fit (ppm) : num [1:720] NA NA 316 317 318 ...  
## $ Seasonally Adjusted CO2 Fit (ppm): num [1:720] NA NA 315 315 315 ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. Year = col\_double(),  
## .. Month = col\_double(),  
## .. `Decimal Date` = col\_double(),  
## .. `Carbon Dioxide (ppm)` = col\_double(),  
## .. `Seasonally Adjusted CO2 (ppm)` = col\_double(),  
## .. `Carbon Dioxide Fit (ppm)` = col\_double(),  
## .. `Seasonally Adjusted CO2 Fit (ppm)` = col\_double()  
## .. )

head(Co2Est)

## # A tibble: 6 x 7  
## Year Month `Decimal Date` `Carbon Dioxide ~ `Seasonally Adj~ `Carbon Dioxide~  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 1958 1 1958. NA NA NA   
## 2 1958 2 1958. NA NA NA   
## 3 1958 3 1958. 316. 314. 316.  
## 4 1958 4 1958. 317. 315. 317.  
## 5 1958 5 1958. 318. 315. 318.  
## 6 1958 6 1958. NA NA 317.  
## # ... with 1 more variable: Seasonally Adjusted CO2 Fit (ppm) <dbl>

summary(Co2Est)

## Year Month Decimal Date Carbon Dioxide (ppm)  
## Min. :1958 Min. : 1.00 Min. :1958 Min. :313.2   
## 1st Qu.:1973 1st Qu.: 3.75 1st Qu.:1973 1st Qu.:328.6   
## Median :1988 Median : 6.50 Median :1988 Median :349.8   
## Mean :1988 Mean : 6.50 Mean :1988 Mean :352.4   
## 3rd Qu.:2002 3rd Qu.: 9.25 3rd Qu.:2003 3rd Qu.:373.2   
## Max. :2017 Max. :12.00 Max. :2018 Max. :407.6   
## NA's :17   
## Seasonally Adjusted CO2 (ppm) Carbon Dioxide Fit (ppm)  
## Min. :314.4 Min. :312.4   
## 1st Qu.:329.0 1st Qu.:328.3   
## Median :349.8 Median :349.4   
## Mean :352.4 Mean :352.1   
## 3rd Qu.:372.9 3rd Qu.:372.8   
## Max. :406.0 Max. :407.3   
## NA's :17 NA's :13   
## Seasonally Adjusted CO2 Fit (ppm)  
## Min. :314.9   
## 1st Qu.:328.4   
## Median :349.3   
## Mean :352.0   
## 3rd Qu.:372.6   
## Max. :405.8   
## NA's :13

Asi que como se pudo observar el dataset nadamas cuenta con datos de fechas y emision de Co2.

## Analisis de Normalidad de los datos (Netflix$Hidden Gem Score)

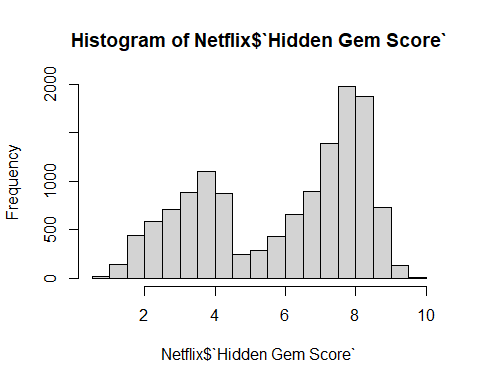
Para el dataset Netflix, determinar si existe normalidad en los datos de la variable Hidden Gem Score.

Modelo = Analisis de normalidad de los datos de Pearson.

Grafico = Histograma.

Hipotesis. H0 = La Variable Hidden Gem Score posee una normalidad en los datos H1 = La Variable Hidden Gem Score NO posee una normalidad en los datos

hist(Netflix$`Hidden Gem Score`)



pearson.test(Netflix$`Hidden Gem Score`)

##   
## Pearson chi-square normality test  
##   
## data: Netflix$`Hidden Gem Score`  
## P = 10286, p-value < 2.2e-16

Pvalue 2.2e-16 es menor a 0.05 por lo tanto se rechaza H0.

Conclusion: debido a que el pvalue es menor a el nivel de significancia rechazamos nuestra H0 por lo tanto podemos determinar que no hay una normalidad en los datos.

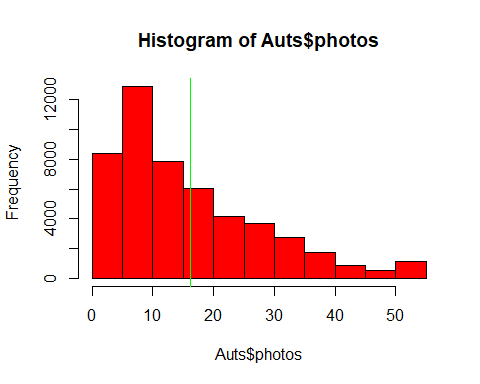
## Analisis de Medias (Auts$photos)

Para el dataset de Autos determinar la media de fotos que las personas suben al vender su auto.

Modelo = Mean, analisis de medias.

Grafico = Histograma.

hist(Auts$photos,col = ("red"))  
mediaphotos <- mean(Auts$photos)  
abline(v=mediaphotos,col="green")



Conclusion: podemos determinar que la media de fotos que las personas suben a la hora de intentar vender sus autos usados es de 16, y que estas rondan entre las 0 y las 55.

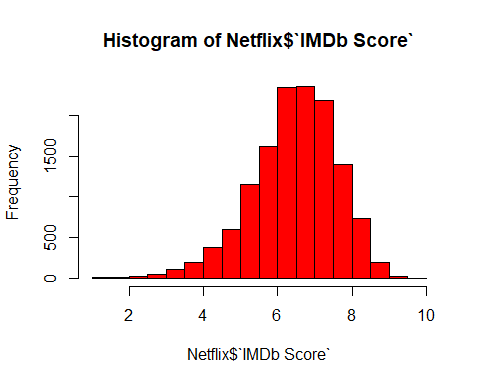
## Analisis de Medias (Auts$photos)

Para el dataset de Netflix determinar la media de Calificacion en IMDB.

Modelo = Mean, analisis de medias.

Grafico = Histograma.

hist(Netflix$`IMDb Score`,col = ("red"))  
mediaIMDB <- mean(Netflix$`IMDb Score`)  
abline(v=mediaIMDB,col="green")



Conclusiones: Podemos determinar que la media de puntuacion del catalogo de netflix en IMDb es de 7, y que a su vez es poco contenido en el catalogo que tiene una puntuacion menor a 3.

## Analisis de Correlacion (Autsyear)

Determinar si existe una relacion entre el Poder de los automoviles y el año de fabricacion.

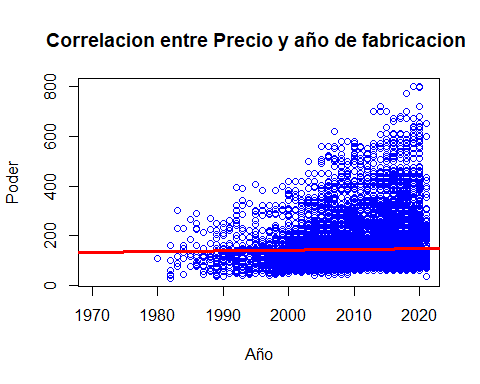
Modelo = Plot.

Grafico = Correlacion.

H0 = No existe relacion entre el Poder y el año de fabricacion del vehiculo.

H1 = Existe una relacion entre el Poder y el año de fabricacion del vehiculo.

plot(Auts$year,Auts$power,col="blue",main="Correlacion entre Precio y año de fabricacion ",xlab="Año",ylab="Poder")  
abline(lm(Auts$power ~ Auts$year), col = 'red', lwd = 3)



cor(Auts$year,Auts$power)

## [1] NA

Indice de Correlacion =Na por lo tanto se acepta la hipotesis nula.

Conclusion: Podemos determinar gracias al indice de correlacion (NA) que no existe una relacion estrecha entre el año de fabricacion y el poder de los vehiculos

## Analisis de Correlacion (Netflix$`Hidden Gem Score` y Netflix$IMDb Score)

Determinar si existe una relacion entre La calificacion interna de netflix y la de IMDb.

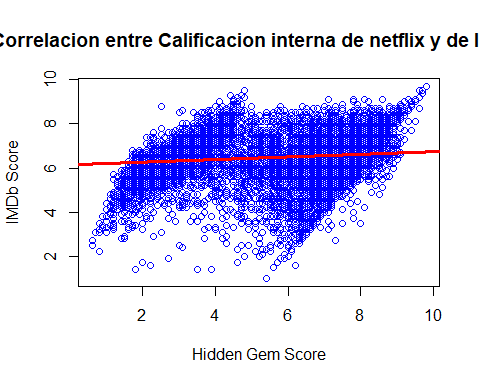
Modelo = Plot.

Grafico = Correlacion.

H0 = No existe relacion entre La calificacion interna de netflix y la de IMDb.

H1 = Existe una relacion entre La calificacion interna de netflix y la de IMDb.

plot(Netflix$`Hidden Gem Score`,Netflix$`IMDb Score`,col="blue",main="Correlacion entre Calificacion interna de netflix y de IMDb ",xlab="Hidden Gem Score",ylab="IMDb Score")  
abline(lm(Netflix$`IMDb Score` ~ Netflix$`Hidden Gem Score`), col = 'red', lwd = 3)



cor(Netflix$`Hidden Gem Score`,Netflix$`IMDb Score`)

## [1] NA

Indice de Correlacion =Na por lo tanto se acepta la hipotesis nula.

Conclusion: Podemos determinar gracias al indice de correlacion (NA) que no existe una relacion entre la calificacion interna que tiene netflix y la que se le ha otorgado en IMDb.

## Analisis de Correlacion (Netflix$`Hidden Gem Score` y Netflix$IMDb Score)

Determinar si existe una relacion entre La calificacion interna de netflix y la de IMDb.

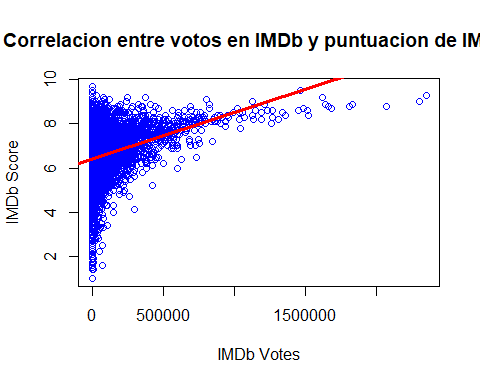
Modelo = Plot.

Grafico = Correlacion.

H0 = No existe relacion entre La calificacion interna de netflix y la de IMDb.

H1 = Existe una relacion entre La calificacion interna de netflix y la de IMDb.

plot(Netflix$`IMDb Votes`,Netflix$`IMDb Score`,col="blue",main="Correlacion entre votos en IMDb y puntuacion de IMDb ",xlab="IMDb Votes",ylab="IMDb Score")  
abline(lm(Netflix$`IMDb Score` ~ Netflix$`IMDb Votes`), col = 'red', lwd = 3)



cor(Netflix$`IMDb Votes`,Netflix$`IMDb Score`)

## [1] NA

Indice de Correlacion =Na por lo tanto se acepta la hipotesis nula.

Conclusion: Podemos determinar que aunque al principio de la grafica pareciese que mientras menos personas hayan votado sera mayor la puntuacion la realidad es que no son variables dependientes.

## Analisis de Prueba de medias con T.test (Autsshift)

determinar tanto grafica como analiticamente si existe una diferencia significativa entre el tipo de cambio y el poder del vehiculo.

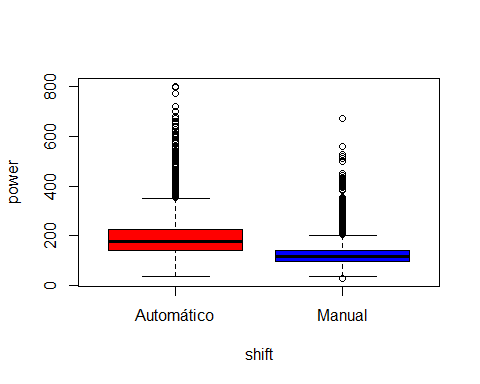
Modelo = T.test.

Grafica = Boxplot.

H0 = no existe una diferencia significativa entre las medias de tipo de cambio y el poder del vehiculo.

H1 = si existe una diferencia significativa entre las medias de tipo de cambio y el poder del vehiculo.

boxplot(power~shift, data=Auts, col=c("red","blue"))



t.test(Auts$power~Auts$shift, alternative="two.sided")

##   
## Welch Two Sample t-test  
##   
## data: Auts$power by Auts$shift  
## t = 96.887, df = 16606, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 75.86044 78.99325  
## sample estimates:  
## mean in group Automático mean in group Manual   
## 197.5716 120.1448

pvalue (2.2e-16) menor a 0.05 se rechaza la H0.

Conclusiones: debido a que el pvalue es menor al nivel de significancia se rechaza la hipotesis nula, por lo tanto podemos determinar que existe una diferencia entre las medias de poder en base al tipo de cambio, teniendo un mayor poder los automoviles Automaticos

## Analisis de Prueba de medias con T.test (Netflix$`Hidden Gem Score` y Netflix$Series or Movie)

determinar tanto grafica como analiticamente si existe una diferencia significativa entre (Movie or Serie) y hidden Gem Score (calificacion interna de netflix).

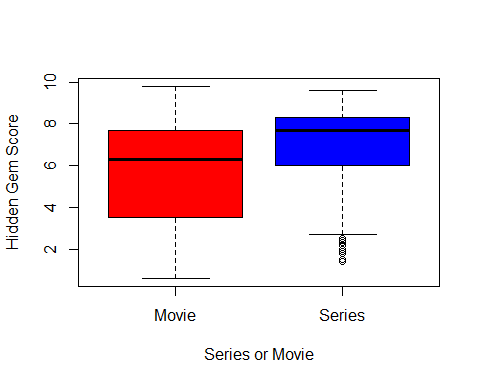
Modelo = T.test.

Grafica = Boxplot.

H0 = no existe una diferencia significativa entre (Movie or Serie) y hidden Gem Score (calificacion interna de netflix).

H1 = si existe una diferencia significativa entre (Movie or Serie) y hidden Gem Score (calificacion interna de netflix).

boxplot(`Hidden Gem Score`~`Series or Movie`, data=Netflix, col=c("red","blue"))



t.test(Netflix$`Hidden Gem Score`~Netflix$`Series or Movie`, alternative="two.sided")

##   
## Welch Two Sample t-test  
##   
## data: Netflix$`Hidden Gem Score` by Netflix$`Series or Movie`  
## t = -34.615, df = 6344.4, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -1.442694 -1.288043  
## sample estimates:  
## mean in group Movie mean in group Series   
## 5.620779 6.986147

pvalue (2.2e-16) menor a 0.05 se rechaza la H0.

Conclusiones: debido a que el pvalue es menor al nivel de significancia se rechaza la hipotesis nula, por lo tanto podemos determinar que existe una diferencia significativa entre las medias de calificacion de las peliculas y las series de netflix siendo las series las mas puntuadas

## Análisis de Prueba de medias con multiples variables (Fuel y kms de Auts)

Examinar tanto de forma grafica como analitica la relacion que existe entre el tipo de combustible y los kms recorridos.

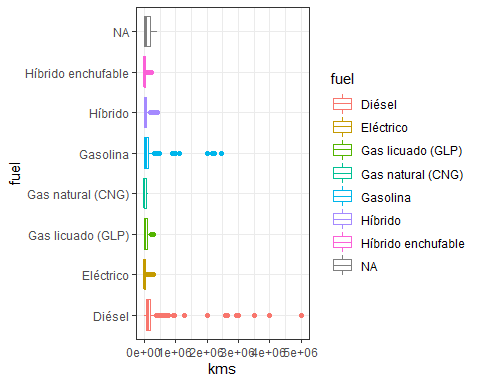
Modelo = Anova.

Grafica = Boxplot.

H0 = No existe relacion entre el tipo de combustible y los km recorridos.

H1 = Existe relacion entre el tipo de combustible y los km recorridos.

ggplot(data= Auts, aes(x=kms, y=fuel,color=fuel ))+geom\_boxplot()+theme\_bw()



anova <- aov(Auts$kms~Auts$fuel)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Auts$fuel 6 4.548e+13 7.581e+12 715.1 <2e-16 \*\*\*  
## Residuals 49947 5.295e+14 1.060e+10   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## 46 observations deleted due to missingness

debido a que el pvalue 2e-16 es menor a 0.05 se rechaza la H0.

Conclusiones: en base a que el pvalue es menor a 0.05 se puede determinar que existe una relacion significativa entre el tipo de combustible y los kilometros recorridos.

## Análisis de Prueba de medias con multiples variables (Runtime e IMDb Score)

Examinar tanto de forma grafica como analitica la relacion que existe entre la duracion de una pelicula o serie y la puntuacion que le dan los usuarios en IMDb.

Modelo = Anova.

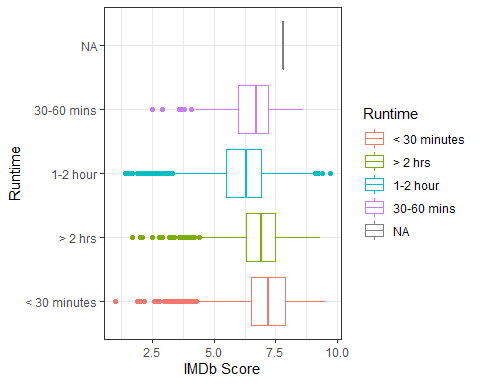
Grafica = Boxplot.

H0 = No existe relacion entre la duracion de una pelicula o serie y la puntuacion que le dan los usuarios en IMDb.

H1 = Existe relacion entre la duracion de una pelicula o serie y la puntuacion que le dan los usuarios en IMDb.

ggplot(data= Netflix, aes(x=`IMDb Score`, y=Runtime,color=Runtime ))+geom\_boxplot()+theme\_bw()

## Warning: Removed 2099 rows containing non-finite values (stat\_boxplot).



anova <- aov(Netflix$`IMDb Score`~Netflix$Runtime)  
summary(anova)

## Df Sum Sq Mean Sq F value Pr(>F)   
## Netflix$Runtime 3 2150 716.7 620.6 <2e-16 \*\*\*  
## Residuals 13376 15448 1.2   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
## 2100 observations deleted due to missingness

debido a que el pvalue 2e-16 es menor a 0.05 se rechaza la H0.

Conclusiones: en base a que el pvalue es menor a 0.05 se puede determinar que existe una relacion significativa Entre el tiempo que dura una pelicula o serie y que tan bien puntuada este, siendo las que duran menos de 30 minutos las mas populares.

## Análisis de ChiCuadrado (Autsfuel)

Examinar tanto grafica como analiticamente la relacion entre el tipo de combustible y si en vendedor es profesional, y determine si existe una diferencia significativa entre las variables

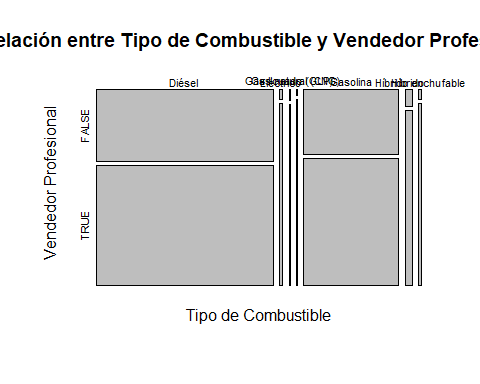
Modelo = Test chi cuadrado.

Grafica = Mosaic plot.

H0 = no existe relacion entre las variables tipo de combustible y vendedor.

H1 = existe relacion entre las variables tipo de combustible y vendedor.

tabla <- table(Auts$fuel,Auts$is\_professional, dnn=c("Tipo de Combustible","Vendedor Profesional"))  
mosaicplot(tabla, main = deparse("Relación entre Tipo de Combustible y Vendedor Profesional"))



tabla

## Vendedor Profesional  
## Tipo de Combustible FALSE TRUE  
## Diésel 11597 19269  
## Eléctrico 29 529  
## Gas licuado (GLP) 7 107  
## Gas natural (CNG) 3 88  
## Gasolina 5659 10980  
## Híbrido 106 1068  
## Híbrido enchufable 27 485

chisq.test(tabla)

##   
## Pearson's Chi-squared test  
##   
## data: tabla  
## X-squared = 944.78, df = 6, p-value < 2.2e-16

el p-value (2.2e-16) es menor a 0.05 por lo cual se rechaza H0.

Conclusiones: podemos determinar debido a que el pvalues es menor a 0.05 que existe una relacion entre el tipo de combustible y si el vendedor es o no profesional

## Análisis de Series de tiempo (Emision de Co2)

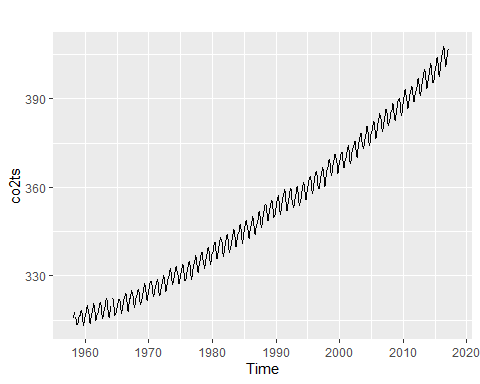
Para el dataset Co2Est realizar un analisis detallado de lines de tiempo en base a la emision de C02

co2ts <- ts(Co2Est$`Carbon Dioxide (ppm)`, start = c(1958,1), frequency = 12)  
print(co2ts)

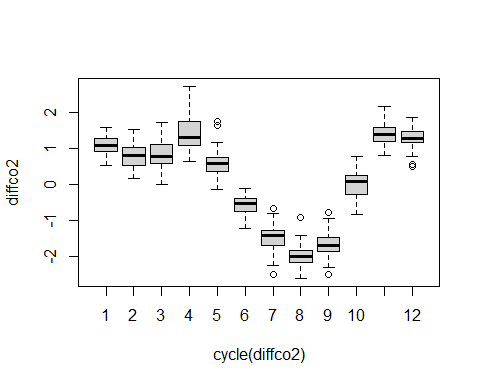
## Jan Feb Mar Apr May Jun Jul Aug Sep Oct  
## 1958 NA NA 315.69 317.45 317.50 NA 315.86 314.93 313.21 NA  
## 1959 315.58 316.48 316.65 317.72 318.29 318.15 316.54 314.80 313.84 313.34  
## 1960 316.43 316.97 317.58 319.03 320.03 319.59 318.18 315.90 314.17 313.83  
## 1961 316.90 317.70 318.54 319.48 320.58 319.77 318.57 316.79 314.99 315.31  
## 1962 317.94 318.55 319.68 320.57 321.01 320.62 319.61 317.40 316.25 315.42  
## 1963 318.74 319.07 319.86 321.39 322.25 321.48 319.74 317.77 316.21 315.99  
## 1964 319.57 NA NA NA 322.26 321.89 320.44 318.69 316.71 316.87  
## 1965 319.44 320.44 320.89 322.14 322.16 321.87 321.21 318.87 317.82 317.30  
## 1966 320.62 321.59 322.39 323.70 324.08 323.75 322.39 320.36 318.64 318.10  
## 1967 322.32 322.50 323.04 324.42 325.00 324.09 322.55 320.92 319.25 319.39  
## 1968 322.57 323.14 323.89 325.02 325.57 325.36 324.14 322.11 320.33 320.25  
## 1969 324.00 324.42 325.63 326.66 327.38 326.70 325.88 323.66 322.38 321.78  
## 1970 325.06 325.98 326.93 328.14 328.08 327.67 326.34 324.69 323.10 323.07  
## 1971 326.17 326.68 327.17 327.79 328.92 328.57 327.36 325.43 323.36 323.57  
## 1972 326.77 327.63 327.75 329.73 330.07 329.09 328.04 326.32 324.84 325.20  
## 1973 328.55 329.56 330.30 331.50 332.48 332.07 330.87 329.31 327.51 327.18  
## 1974 329.35 330.71 331.48 332.65 333.09 332.25 331.18 329.39 327.43 327.37  
## 1975 330.40 331.40 332.04 333.31 333.97 333.60 331.90 330.06 328.56 328.34  
## 1976 331.75 332.56 333.50 334.58 334.88 334.34 333.05 330.94 329.30 328.94  
## 1977 332.93 333.42 334.70 336.07 336.75 336.27 334.92 332.75 331.59 331.16  
## 1978 334.97 335.38 336.64 337.76 338.01 337.89 336.54 334.68 332.76 332.55  
## 1979 336.23 336.76 337.96 338.89 339.47 339.29 337.73 336.08 333.92 333.86  
## 1980 338.01 338.36 340.07 340.77 341.47 341.17 339.56 337.60 335.88 336.02  
## 1981 339.24 340.48 341.38 342.51 342.91 342.25 340.49 338.43 336.69 336.86  
## 1982 340.75 341.61 342.70 343.57 344.14 343.35 342.06 339.81 337.98 337.86  
## 1983 341.38 342.52 343.10 344.94 345.76 345.32 343.98 342.38 339.87 339.99  
## 1984 343.70 344.50 345.28 347.07 347.43 346.80 345.39 343.28 341.07 341.35  
## 1985 344.97 345.99 347.43 348.35 348.93 348.25 346.56 344.68 343.09 342.80  
## 1986 346.30 346.95 347.85 349.55 350.21 349.55 347.94 345.90 344.85 344.17  
## 1987 348.02 348.47 349.41 350.99 351.85 351.26 349.51 348.10 346.45 346.36  
## 1988 350.43 351.73 352.22 353.59 354.22 353.79 352.38 350.43 348.73 348.88  
## 1989 352.76 353.07 353.68 355.42 355.67 355.12 353.90 351.67 349.80 349.99  
## 1990 353.66 354.70 355.38 356.20 357.16 356.23 354.81 352.91 350.96 351.18  
## 1991 354.72 355.75 357.16 358.60 359.34 358.24 356.17 354.02 352.15 352.21  
## 1992 355.99 356.72 357.81 359.15 359.66 359.25 357.02 355.00 353.01 353.31  
## 1993 356.70 357.16 358.38 359.46 360.28 359.60 357.57 355.52 353.69 353.99  
## 1994 358.37 358.91 359.97 361.27 361.69 360.94 359.55 357.48 355.84 356.00  
## 1995 359.97 361.00 361.64 363.45 363.80 363.26 361.90 359.45 358.05 357.76  
## 1996 362.05 363.24 364.02 364.72 365.41 364.97 363.65 361.48 359.45 359.61  
## 1997 363.18 363.99 364.56 366.36 366.80 365.63 364.47 362.50 360.19 360.77  
## 1998 365.33 366.15 367.31 368.61 369.30 368.88 367.64 365.77 363.90 364.24  
## 1999 368.15 368.87 369.59 371.14 371.00 370.35 369.27 366.93 364.64 365.13  
## 2000 369.14 369.46 370.51 371.66 371.83 371.69 370.12 368.12 366.62 366.73  
## 2001 370.28 371.50 372.12 372.87 374.02 373.31 371.62 369.55 367.96 368.10  
## 2002 372.44 373.08 373.52 374.86 375.55 375.40 374.02 371.48 370.70 370.25  
## 2003 374.68 375.62 376.11 377.65 378.35 378.13 376.61 374.48 372.98 373.00  
## 2004 376.79 377.36 378.39 380.50 380.61 379.55 377.76 375.84 374.05 374.22  
## 2005 378.34 379.61 380.17 382.05 382.24 382.08 380.67 378.67 376.42 376.80  
## 2006 381.37 382.02 382.56 384.37 384.92 384.03 382.28 380.48 378.81 379.06  
## 2007 382.58 383.71 384.34 386.23 386.41 385.87 384.45 381.84 380.86 380.86  
## 2008 385.07 385.84 385.83 386.77 388.51 388.05 386.25 384.08 383.09 382.78  
## 2009 386.65 387.12 388.51 389.57 390.16 389.62 388.07 386.08 384.65 384.33  
## 2010 388.55 390.07 391.01 392.38 393.22 392.24 390.33 388.52 386.84 387.16  
## 2011 391.30 391.92 392.45 393.37 394.28 393.69 392.60 390.21 389.00 388.93  
## 2012 393.07 393.35 394.36 396.43 396.87 395.88 394.52 392.54 391.13 391.01  
## 2013 395.61 396.85 397.26 398.35 399.98 398.87 397.37 395.41 393.39 393.70  
## 2014 397.93 398.10 399.47 401.33 401.88 401.31 399.07 397.21 395.40 395.65  
## 2015 399.85 400.31 401.51 403.45 404.11 402.88 401.60 399.00 397.50 398.28  
## 2016 402.65 404.16 404.86 407.57 407.65 407.00 404.50 402.24 401.01 401.50  
## 2017 406.07 406.64 NA NA NA NA NA NA NA NA  
## Nov Dec  
## 1958 313.33 314.67  
## 1959 314.81 315.58  
## 1960 315.00 316.19  
## 1961 316.10 317.01  
## 1962 316.69 317.70  
## 1963 317.06 318.35  
## 1964 317.68 318.71  
## 1965 318.87 319.42  
## 1966 319.79 321.03  
## 1967 320.72 321.96  
## 1968 321.32 322.89  
## 1969 322.85 324.12  
## 1970 324.01 325.13  
## 1971 324.80 326.01  
## 1972 326.50 327.55  
## 1973 328.16 328.64  
## 1974 328.46 329.57  
## 1975 329.49 330.76  
## 1976 330.31 331.68  
## 1977 332.40 333.85  
## 1978 333.92 334.95  
## 1979 335.29 336.73  
## 1980 337.10 338.21  
## 1981 338.36 339.61  
## 1982 339.26 340.49  
## 1983 341.15 342.99  
## 1984 342.98 344.22  
## 1985 344.24 345.56  
## 1986 345.66 346.90  
## 1987 347.81 348.96  
## 1988 350.07 351.34  
## 1989 351.29 352.52  
## 1990 352.83 354.21  
## 1991 353.75 354.99  
## 1992 354.16 355.40  
## 1993 355.33 356.80  
## 1994 357.58 359.04  
## 1995 359.56 360.70  
## 1996 360.76 362.33  
## 1997 362.43 364.28  
## 1998 365.46 366.97  
## 1999 366.68 368.01  
## 2000 368.29 369.53  
## 2001 369.68 371.24  
## 2002 372.08 373.78  
## 2003 374.34 375.69  
## 2004 375.84 377.44  
## 2005 378.31 379.96  
## 2006 380.14 381.66  
## 2007 382.35 383.61  
## 2008 384.01 385.11  
## 2009 386.05 387.49  
## 2010 388.67 389.81  
## 2011 390.24 391.80  
## 2012 392.95 394.34  
## 2013 395.19 396.82  
## 2014 397.22 398.79  
## 2015 400.24 401.89  
## 2016 403.64 404.55  
## 2017 NA NA

autoplot(co2ts, ts.colour = "blue", ts.linetype = "dashed")

## Warning: Ignoring unknown parameters: ts.colour, ts.linetype

 En base al analisis de tiempo general podemos determinar que el aumento de Co2 no es estacionario, sino que sigue en incremento constante en base al tiempo transcurrido

diffco2<-diff(co2ts)  
boxplot(diffco2~cycle(diffco2))

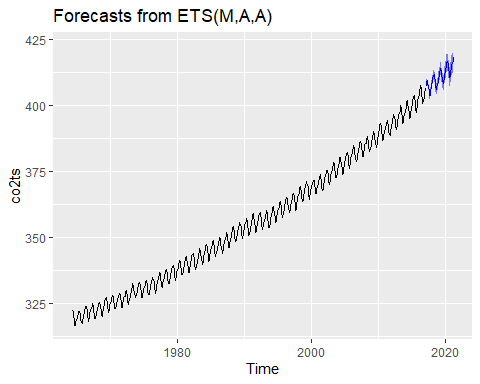


Gracias a la funcion Diffco2 podemos observar en que meses del año sucede una mayor emision de Co2 los cuales son Enero, Abril, Noviembre y Diciembre (1,4,11 y 12) mientras que en julio, agosto y septiembre la emision de estos baja.

forecast1<-forecast(co2ts, level = c(95), h = 50)

## Warning in ets(object, lambda = lambda, biasadj = biasadj,  
## allow.multiplicative.trend = allow.multiplicative.trend, : Missing values  
## encountered. Using longest contiguous portion of time series

autoplot(forecast1)



Y gracias al analisis con la herramienta de forecast podemos predecir que para los años siguientes a la toma de datos se ha incrementado de manera constante la emision de Co2