Week 4

Tomás

22 de noviembre de 2019

# Editing text variables

#Video 1  
  
data <- read.csv("Baltimore\_Fixed\_Speed\_Cameras.csv")  
names(data)

## [1] "address" "direction" "street" "crossStreet"   
## [5] "intersection" "Location.1"

tolower(names(data)) # Deja todo en minuscula (se puede en mayuscula)

## [1] "address" "direction" "street" "crossstreet"   
## [5] "intersection" "location.1"

splitnames <- strsplit(names(data), "\\.") #splitea un string  
splitnames[[6]][1]

## [1] "Location"

review <- data.frame( hola\_chao = "hola", chao\_hola = "chao")  
names(review)# Quiero quitar los "\_"

## [1] "hola\_chao" "chao\_hola"

sub("\_","",names(review)) #Quita los "\_" y los cambia por nada "" (gsub() reemplaza todos y no solo el primero)

## [1] "holachao" "chaohola"

grep("Alameda", data$intersection) # Busca las coincidencias y me entrega los indices

## [1] 65 69 79

grep("Alameda", data$intersection, value = TRUE) # Busca las coincidencias y me entrega los valores

## [1] "E 33rd & The Alameda" "The Alameda & 33rd St"   
## [3] "Harford \n & The Alameda"

table(grepl("Alameda", data$intersection)) # Entre el numero de veces que aparece "Alameda"

##   
## FALSE TRUE   
## 77 3

library (stringr)  
nchar("Jeffrey Leek") #Me da la cantidad de caracteres

## [1] 12

substr("Jefrey Leek",1,7) #Me da los caracteresd del 1 al 7

## [1] "Jefrey "

paste("hola, Chao") #Pega strings

## [1] "hola, Chao"

paste0("hola","chao") #Pega sin espacios

## [1] "holachao"

str\_trim("Tomas ") #Quita los espacios innecesarios

## [1] "Tomas"

# Video 2

## Metacaracteres

Al buscar caracteres dentro de un texto puedo querer buscarlos de forma mas especifica que simplemente por la palabra (Funcion grep()):

* ^Hola: Me busca todas las lineas que comienzen con Hola (Si esta en medio no lo identifica)
* Hola$: Busca todas las frases que terminen con Hola
* [Bb][Uu][Ss][Hh]: Busca la palabra “Bush” en todos sus formatos, independiente si es mayuscula o minuscula
* [[1]](#footnote-1)[a-zA-Z]: Busca linea que comienzen con un numero del 0-9 y sigan con letras de la “a” a la “z”
* [^?.]$: Busca linea que no terminen en “?” ni “.”
* 9.11: Busca cualquier frase que tenga un 9 seguido de cualquier caracter y despues un 11
* flood|fire: Busca cualquiera de las dos palabras (pueden ser infinitos |)
* ^([Gg]ood|[Bb]ad): Busca la palabra good o bad en un principio (El parentesis hace que “^” sea para ambos)
* ^Hola (.)?: Me busca la palabra Hola en el inicio y como alternativa el punto (Se pone el slash debido a que el punto simboliza cualquier metacaracter pero quiero que me busque el punto
* (.\*): Me busca cualquier cosas dentro de un parentesis repetido las veces que quiera
* [Bb](+(%5E%20)+%20+){1,5} debate: Busca la palabra bush, despues un espacio, seguido de algo que no es un espacio, otro espacio y la palabra debate que este separado por 1 a 5 veces esa combinacion

# Video 3

date <- date()  
class(date)

## [1] "character"

date1 <- Sys.Date()  
class(date1)

## [1] "Date"

format(date1, "%A %B %y") # Convierte la fecha a un formato que el programa entiende

## [1] "viernes noviembre 19"

x <- c("1junio1960","2junio1960") ; x

## [1] "1junio1960" "2junio1960"

y <- as.Date(x,"%d%b%Y") ; y

## [1] "1960-06-01" "1960-06-02"

y[1] - y[2]

## Time difference of -1 days

weekdays(date1) # Dia

## [1] "viernes"

months(date1) #Mes

## [1] "noviembre"

julian(date1) #Me entrega los dias desde el "origen"

## [1] 18222  
## attr(,"origin")  
## [1] "1970-01-01"

library(lubridate)

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

#Solo sirve para pasar numeros a fechas  
ymd("20140108") #Convierte a fecha en el orden que le entrego (year,month,day)

## [1] "2014-01-08"

myd("11/2014.11") #(month,year,day), Le puedo entregar casi cualquier cosa entre medio

## [1] "2014-11-11"

ymd\_hms("20140108 10:15:03") #Puedo agregarle la hora

## [1] "2014-01-08 10:15:03 UTC"

ymd\_hms("20140108 10:15:03", tz="Pacific/Auckland") #Puedo agregarle la hora y la zona horaria

## [1] "2014-01-08 10:15:03 NZDT"

1. 0-9 [↑](#footnote-ref-1)