

Covid 19 en México

Murpholinox Peligro

28 04 2020

```
# Carga los paquetes necesarios.
library(ggplot2)
library(ggdark)
library(latex2exp)
library(lubridate)
library(dplyr)
# Configura el directorio de trabajo.
setwd("/home/murphy/Repos/plotcovid19mx")
```

```
# Descarga datos del European CDC.
wget -O full.csv https://opendata.ecdc.europa.eu/covid19/casedistribution/csv
# Obtiene las líneas correspondientes a México.
grep Mex full.csv > mex.csv
# Solo ocupamos la columna 1 y 6 (fecha y decesos por día).
awk -F "," '{print $1,"" $6}' mex.csv > clean.csv
# Añade nombre a las columnas.
echo "fecha,decesos" >> clean.csv
# Revierte los datos en la lista.
tac clean.csv > clean_r.csv
```

```
## --2020-04-28 20:04:12-- https://opendata.ecdc.europa.eu/covid19/casedistribution/csv
## Resolving opendata.ecdc.europa.eu (opendata.ecdc.europa.eu)... 212.181.0.63
## Connecting to opendata.ecdc.europa.eu (opendata.ecdc.europa.eu)|212.181.0.63|:443... connected.
## HTTP request sent, awaiting response... 301 Moved Permanently
## Location: https://opendata.ecdc.europa.eu/covid19/casedistribution/csv/ [following]
## --2020-04-28 20:04:18-- https://opendata.ecdc.europa.eu/covid19/casedistribution/csv/
## Reusing existing connection to opendata.ecdc.europa.eu:443.
## HTTP request sent, awaiting response... 200 OK
## Length: 792917 (774K) [application/octet-stream]
## Saving to: 'full.csv'
##
##      OK ..... 6% 129K 6s
##     50K ..... 12% 258K 4s
##    100K ..... 19% 262K 3s
##    150K ..... 25% 260K 3s
##    200K ..... 32% 253K 2s
##    250K ..... 38% 258K 2s
##    300K ..... 45% 1.19M 2s
##    350K ..... 51% 323K 1s
##    400K ..... 58% 267K 1s
##    450K ..... 64% 260K 1s
##    500K ..... 71% 1.09M 1s
##    550K ..... 77% 242K 1s
```

```
##      600K ..... 83% 333K 0s
##      650K ..... 90% 266K 0s
##      700K ..... 96% 261K 0s
##      750K ..... 100% 24.9M=2.7s
```

```
##
```

```
## 2020-04-28 20:04:21 (286 KB/s) - 'full.csv' saved [792917/792917]
```

```
# Carga los datos limpios a R.
```

```
datos <- read.csv("~/Repos/plotcovid19mx/clean_r.csv")
```

```
# Cambia el formato de la fecha de d/m/y a y-m-d.
```

```
datos$newdate <- lubridate::dmy(datos$fecha)
```

```
# Crea una nueva variable con un nuevo formato para la fecha (a números enteros).
```

```
# Esto se hace porque x con formato de fecha impide que el algoritmo para
```

```
# obtener el modelo exponencial llegue a una solución.
```

```
xmax <- max(length(datos$fecha))
```

```
datos$number <- seq(0,xmax-1)
```

```
# Necesitamos los días del brote en México (después del 20-marzo).
```

```
smalldf<-datos %>%
```

```
  filter(number >= 72)
```

```
# Ordena los datos a usar en un tibble.
```

```
x<-smalldf$number
```

```
x<-x-71
```

```
y<-smalldf$decesos
```

```
y2<-y+1
```

```
nice<-tibble(x,y)
```

```
dummy<-tibble(x,y2)
```

```
# Guarda datos finales en formato CSV.
```

```
write.csv(nice, file=~"/Repos/plotcovid19mx/nice.csv")
```

```
# Crea una gráfica base.
```

```
p <- ggplot(data = nice, aes(x=x, y=y)) + geom_point()
```

```
# Crea modelo lineal.
```

```
linm <- lm(log(y2)~x, data = dummy)
```

```
# Obtiene los parámetros del modelo exponencial.
```

```
a1 <- exp(coef(linm)[1])
```

```
b1 <- coef(linm)[2]
```

```
# Aplica los parámetros del modelo exponencial en la creación del mismo.
```

```
expm <- nls(y ~ a * exp(b * x), start = list(a=a1, b=b1), data = nice)
```

```
summary(expm)
```

```
##
```

```
## Formula: y ~ a * exp(b * x)
```

```
##
```

```
## Parameters:
```

```
##      Estimate Std. Error t value Pr(>|t|)
```

```
## a   5.88780      2.50218   2.353  0.0239 *
```

```
## b   0.07408      0.01234   6.005  5.6e-07 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## Residual standard error: 24.29 on 38 degrees of freedom
```

```
##
```

```
## Number of iterations to convergence: 10
```

```
## Achieved convergence tolerance: 3.809e-06
```

```

a2 <- coef(expm)[1]
b2 <- coef(expm)[2]

# Agrega el ajuste exponencial a la gráfica base,
p2 <- p +
  stat_smooth(method = 'nls', formula = y ~ a * exp(b * x), se=FALSE,
    method.args = list(start = list(a = a2, b = b2))) +
  # la ecuación de la exponencial,
  annotate("label", x = 10, y = 100,
    label = TeX(sprintf("$y = %.2f e^{%.2fx}$", a2, b2))) +
  # los títulos necesarios,
  ylab("Decesos") + xlab("Día") +
  ggtitle("Decesos por Covid-19 en México, a partir del 20 de marzo.")
p2

```

Decesos por Covid-19 en México, a partir del 20 de marzo.

