

# Covid 19 Mexico

Murpholinox Peligro

13 de abril 2020

```
# Carga paquetes
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 líneas correspondientes a México
grep Mex full.csv > mex.csv
# Solo ocupamos la columna 1 y 6 (fecha, decesos por día)
awk -F "," '{print $1," " $6}' mex.csv > clean.csv
# Añade nombre a las columnas
echo "fecha,decesos" >> clean.csv
# Ahora ordena los datos
tac clean.csv > clean_r.csv
```

```
## --2020-04-14 15:27:11-- 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]
## --2020-04-14 15:27:12-- 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: 551354 (538K) [application/octet-stream]
## Saving to: 'full.csv'
##
##      OK ..... 9% 136K 4s
##    50K ..... 18% 272K 2s
##   100K ..... 27% 273K 2s
##   150K ..... 37% 138K 2s
##   200K ..... 46% 3.21M 1s
##   250K ..... 55% 291K 1s
##   300K ..... 65% 270K 1s
##   350K ..... 74% 273K 1s
##   400K ..... 83% 1.09M 0s
##   450K ..... 92% 335K 0s
##   500K ..... 100% 230K=2.0s
##
```

```
## 2020-04-14 15:27:14 (267 KB/s) - 'full.csv' saved [551354/551354]
```

```
# Carga los datos semi limpios
datos <- read.csv("~/Repos/plotcovid19mx/clean_r.csv")
# Cambia el formato de la fecha
datos$newdate <- lubridate::dmy(datos$fecha)

# Ahora crea una nueva variable con nuevo formato para la fecha
xmax <- max(length(datos$fecha))
datos$number <- seq(0,xmax-1)
# Necesitamos los días del outbreak en México (después del 20-marzo)
smalldf<-datos %>%
  filter(number >= 72)
# Ordena los datos
x<-smalldf$number
x<-x-71
y<-smalldf$decesos
nice<-tibble(x,y)
# Guarda datos finales
write.csv(nice, file=~ /Repos/plotcovid19mx/nice.csv")
# Crea una gráfica base
p <- ggplot(data = nice, aes(x=x, y=y)) + geom_point()

# Crea el modelo exponencial
m <-nls(y~a*exp(b*x), start = list(a=0.01, b=0.15))
# Imprime información del modelo
summary(m)
```

```
##
## Formula: y ~ a * exp(b * x)
##
## Parameters:
##   Estimate Std. Error t value Pr(>|t|)
## a   1.91634    0.72222   2.653   0.0139 *
## b   0.11739    0.01664   7.054 2.71e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.993 on 24 degrees of freedom
##
## Number of iterations to convergence: 9
## Achieved convergence tolerance: 6.174e-06

# Agrega el ajuste con los parámetros del modelo
p2 <- p +
  stat_smooth(method = 'nls', formula = y ~ a * exp(b * x), se=FALSE,
    method.args = list(start = list(a = 1.7, b = 0.12))) +
  # la ecuación de la exponencial,
  annotate("label", x=5, y=30, label=TeX('$y = 1.7 e^{0.12 x}$')) +
  # los títulos necesarios,
  ylab("Decesos") + xlab("Día") + ggtitle("Decesos por covid-19 (20-03/13-04)")
# y cambia el tema base dependiendo del formato de salida
if (knitr::is_html_output()) {
  p2 + dark_theme_gray(base_size = 15)
} else if (knitr::is_latex_output()) {
  p2 + theme_light(base_size = 15)
```

}

## Decesos por covid-19 (20-03/13-04)

