



## Proyecto Final

Emiliano Villicaña García  
UNAM FE

Eduardo Aguilar Vázquez  
UNAM FE

---

*Keywords:* "El gran economista debe poseer una rara combinación de dotes. Debe ser matemático, historiador, estadista y filósofo. Debe comprender los símbolos y hablar con palabras corrientes. Debe contemplar lo particular en términos de lo general y tocar lo abstracto y lo concreto con el mismo vuelo del pensamiento. Debe estudiar el presente a la luz del pasado y con vistas al futuro. Ninguna parte de la naturaleza del hombre o de sus instituciones debe quedar por completo fuera de su consideración. Debe estar tan fuera de la realidad como un artista y tan cerca de la tierra como un político." J.M. Keynes .

---

### 1. Instrucciones

##El proyecto constará de dos partes con una duración de una semana respectivamente. Para la primera parte deberán elegir una pregunta, misma que se responderá a través de los datos de instancias oficiales. Así como crear una gráfica sencilla que represente la pregunta que eligieron. Para la segunda parte, harán la presentación de su pantalla y nos describirán su código. Veamos las instrucciones de la Parte 1:

**1.1. Deberás responder una pregunta de tu interés, la cual deberá sustentarse con datos de instancias oficiales. Algunos ejemplos son:**

- ¿Cuánto creció el desempleo?
- ¿Cuál fue el número de contagiados de COVID-19 en el mes de Marzo?
- ¿Qué producto se vende más en el negocio de mis papás?
- ¿Cuál es el país con menor brecha de género?

### 1.2. Necesitarás que existan datos e importarlos.

Recuerden la importancia del manejo de datos, siempre los vamos a necesitar para trabajar, es nuestra herramienta para darle vida a la teoría, por ello es recomendable que sepan dónde encontrar rápidamente la información y que tenga un formato sencillo (.xlsx, .csv, .txt) o de la API de Banco Mundial.

### 1.3. Describir sus datos y hacer transformaciones

Una vez que tengan los datos deberán describirlos contestando las siguientes preguntas:

- ¿Cuántos datos tengo?
- ¿Qué tipo de variables son?
- ¿Están en el formato correcto?
- ¿Tengo errores de captura o datos imposibles?

Dependiendo de su pregunta necesitarán crear variables, agrupar, filtrar, obtener estadística descriptiva, etc.

### 1.4. Realicen una gráfica sencilla

Dependiendo del tipo de relación o variables que tengan, deberán elegir la gráfica que mejor represente su pregunta.

Información que ya sabemos para validar el procedimiento. ¿Cuántos confirmados tenemos de CO

```
R> library(tidyverse)
R> library(covidMex)
```

```
R> df <- covidOfficialMx()
```

```
Warning in covidMex::GetFromSSA(date = date, neat = neat): Please keep in mind
that the official report on Covid-19 cases has not a version control system or
whatsoever, therefore it's still difficult to match a specific date to a version
of the report. The latest version available will be downloaded (it can be from
yesterday's).
```

```
R> df %>%
R+   filter(RESULTADO == 'Positivo a SARS-CoV-2') %>% ## Esto debería ser una etiqueta com
R+   tally()
```

```
# A tibble: 1 x 1
```

```
      n
  <int>
1  71105
```

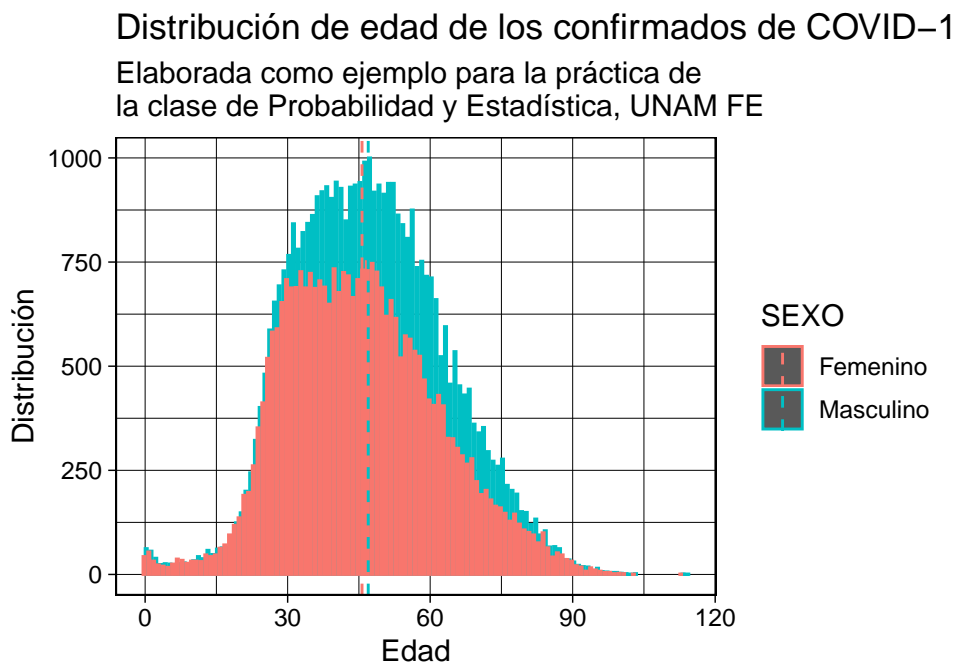
Como el dato concuerda podemos adentrarnos más. De los confirmados ¿Cuántos son hombres y cuántos son mujeres?

```
R> df %>%
R+   filter(RESULTADO == 'Positivo a SARS-CoV-2') %>% ## De verdad, esto debería ser una e
R+   group_by(SEXO) %>%
R+   tally()

# A tibble: 2 x 2
  SEXO      n
  <fct>  <int>
1 Femenino 30642
2 Masculino 40463
```

De los confirmados por sexo ¿Cuál es la distribución de edad?

```
R> confirmados <- df %>%
R+   filter(RESULTADO == 'Positivo a SARS-CoV-2')
R> mu <- confirmados %>% group_by(SEXO) %>% summarise(media = mean(EDAD))
R> ggplot(confirmados,aes(x = EDAD, col = SEXO)) + geom_bar(stat = 'count', position = 'do
R+   geom_vline(data = mu, aes(xintercept=media, color=SEXO),
R+               linetype="dashed") + ## Ahora la hacemos bonita
R+   theme_linedraw() + labs(
R+     title = "Distribución de edad de los confirmados de COVID-19 por sexo",
R+     subtitle = 'Elaborada como ejemplo para la práctica de \nla clase de Probabilidad y
R+     caption = 'Elaboración propia con datos de la Secretaría de Salud',
R+     x = 'Edad',
R+     y = 'Distribución' )
```



Elaboración propia con datos de la Secretaría de Salud

**Affiliation:**