

# ICO 187 ANÁLISIS DE DATOS

## CLASE 29: ANÁLISIS DE DATOS CON R

Año: 2021  
Profesor: Sebastián Egaña

### 1. Transform

Recordemos como generamos la edad en su momento:

```
data <- mutate(data, fecha_nacimiento_2 = as.POSIXct(fecha_nacimiento, format = "%d-%m-%Y"))
```

Lo que utilizando el operador pipe

```
library(dplyr)

data <- data %>%
  mutate(fecha_nacimiento_2 = as.POSIXct(fecha_nacimiento, format = "%d-%m-%Y"))
```

Veamos como generar una variable relacionada con la edad:

```
data <- data %>%
  mutate(hoy = as.POSIXct("2021-04-25", format = "%Y-%m-%d"),
         edad = (hoy - fecha_nacimiento_2)/365)
```

### 2. Visualize

Crearemos un gráfico simple, utilizando la librería ggplot2<sup>1</sup>.

Revisemos el Cheatsheet asociado:

[Enlace acá](#)

En relación a esto, un gráfico utilizando ggplot2 se construye a través de los siguientes elementos:

1. Un set de datos.
2. Un sistema de coordenadas.
3. Figuras geométricas (geoms) o representaciones visuales de los datos.

Para cambiar la forma en que se muestran los datos se pueden editar las figuras geométricas (geoms), cambiando color, tamaño y la ubicación de los puntos en x y/o en y.

Inclusive se plantea una plantilla para esto:

---

<sup>1</sup>Recordar instalar y cargar la librería

```
ggplot (data = <DATA> ) +  
  <GEOM_FUNCTION> (mapping = aes( <MAPPINGS> ),  
  stat = <STAT> , position = <POSITION> ) +  
  <COORDINATE_FUNCTION> +  
  <FACET_FUNCTION> +  
  <SCALE_FUNCTION> +  
  <THEME_FUNCTION>
```

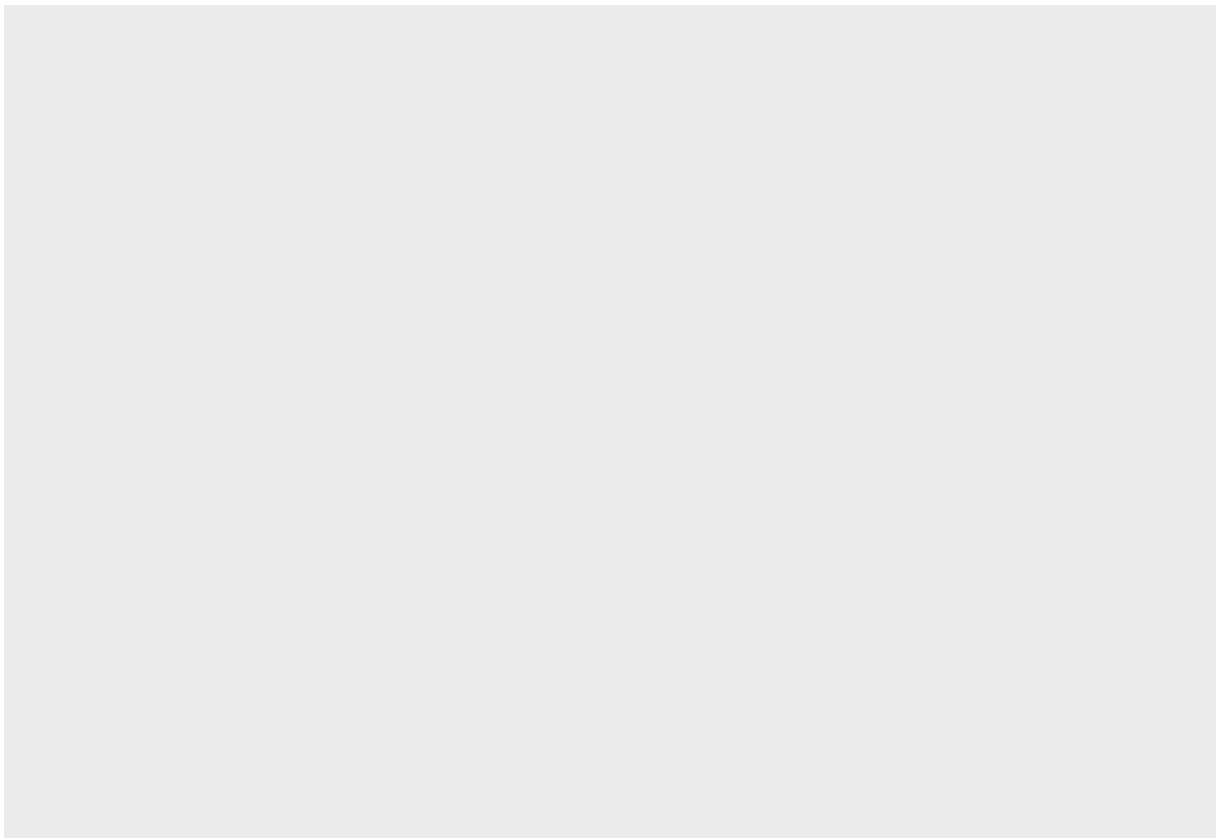
En donde las dos primeras líneas son necesarias.

Apliquemos esto a los datos importados. Referencia relevante para esta parte [Enlace acá](#)

1. Asignamos el set de datos:

```
p <- ggplot(data)
```

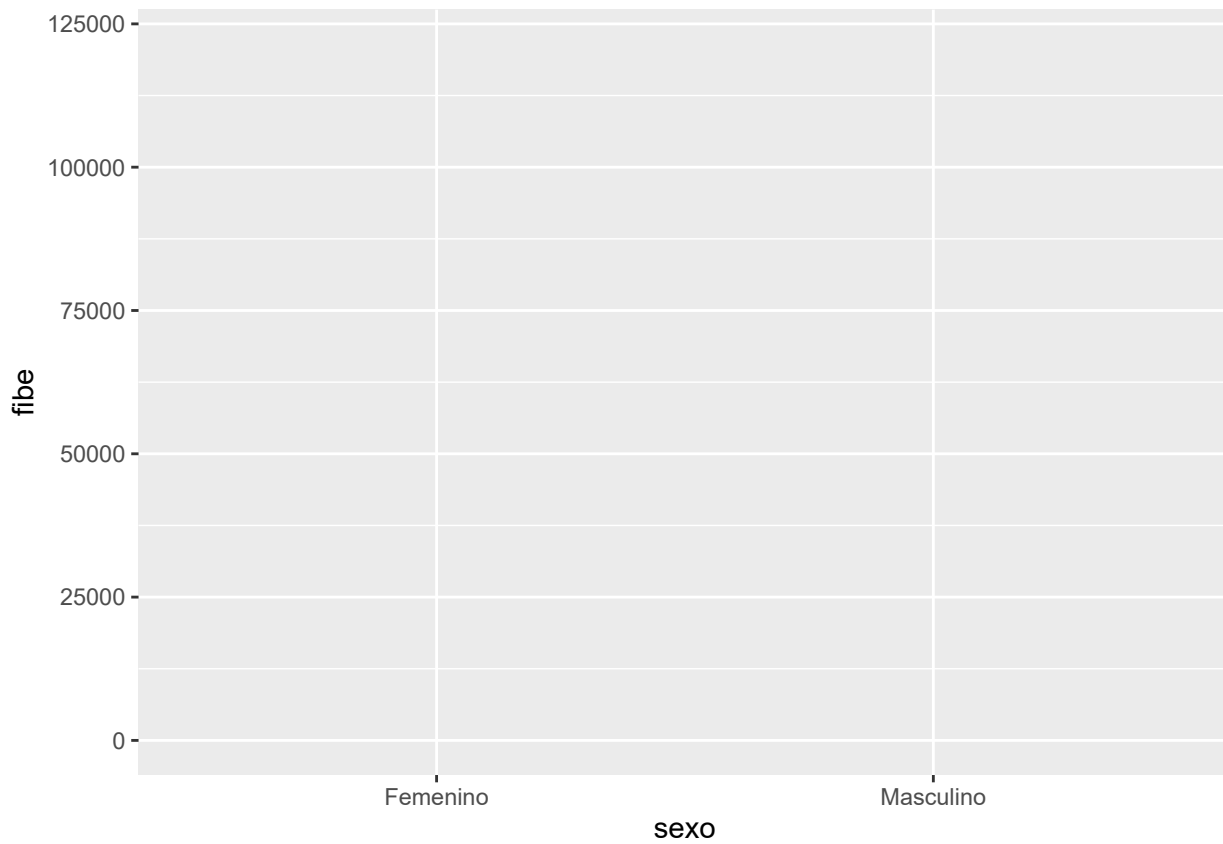
```
p
```



2. Ahora declaramos el sistema de coordenadas:

```
p <- ggplot(data) + aes(x = sexo, y = fibe)
```

```
p
```

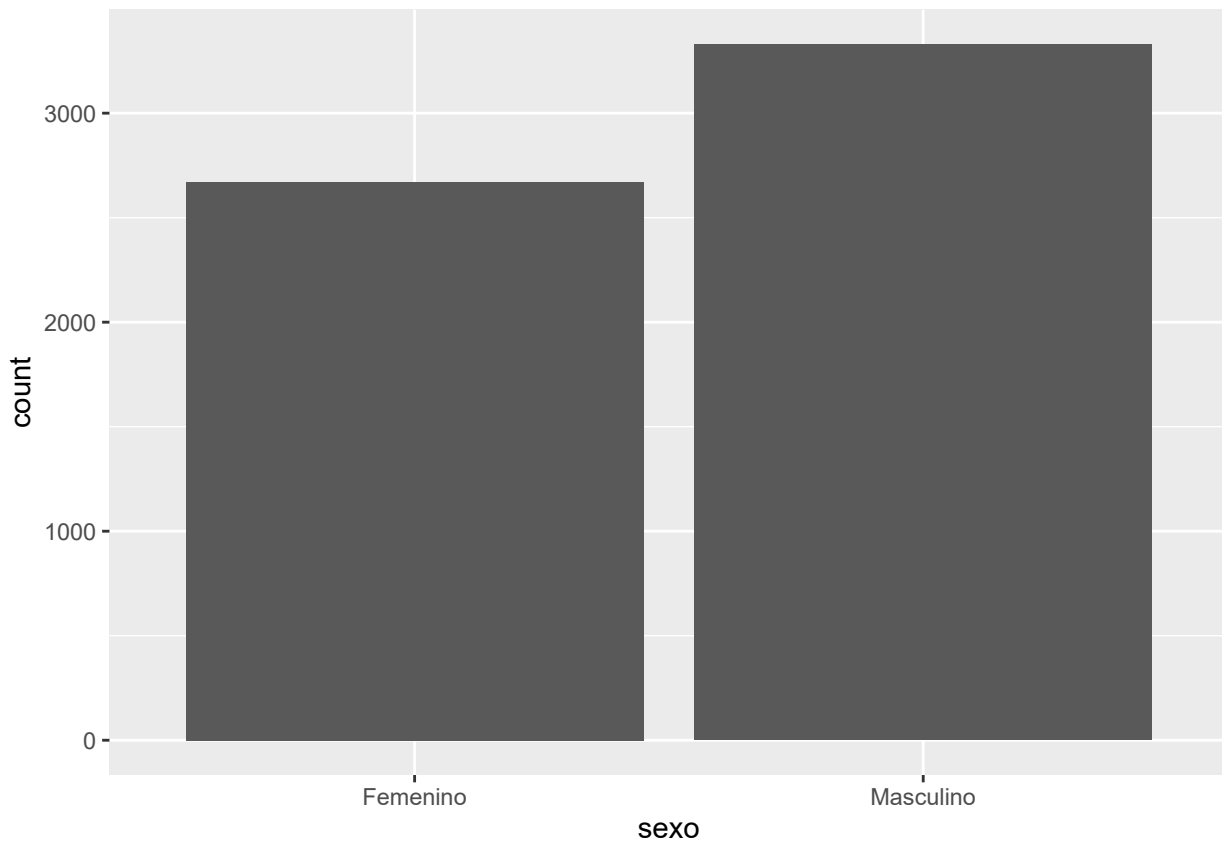


- Cabe mencionar, que cada elemento que se agrega dentro de un gráfico creado utilizando con ggplot, se realiza utilizando el signo “+” para esto.

3. Declaremos ahora el tipo de figura que graficaremos:

```
p <- ggplot(data) + aes(x = sexo) + geom_bar()
```

```
p
```



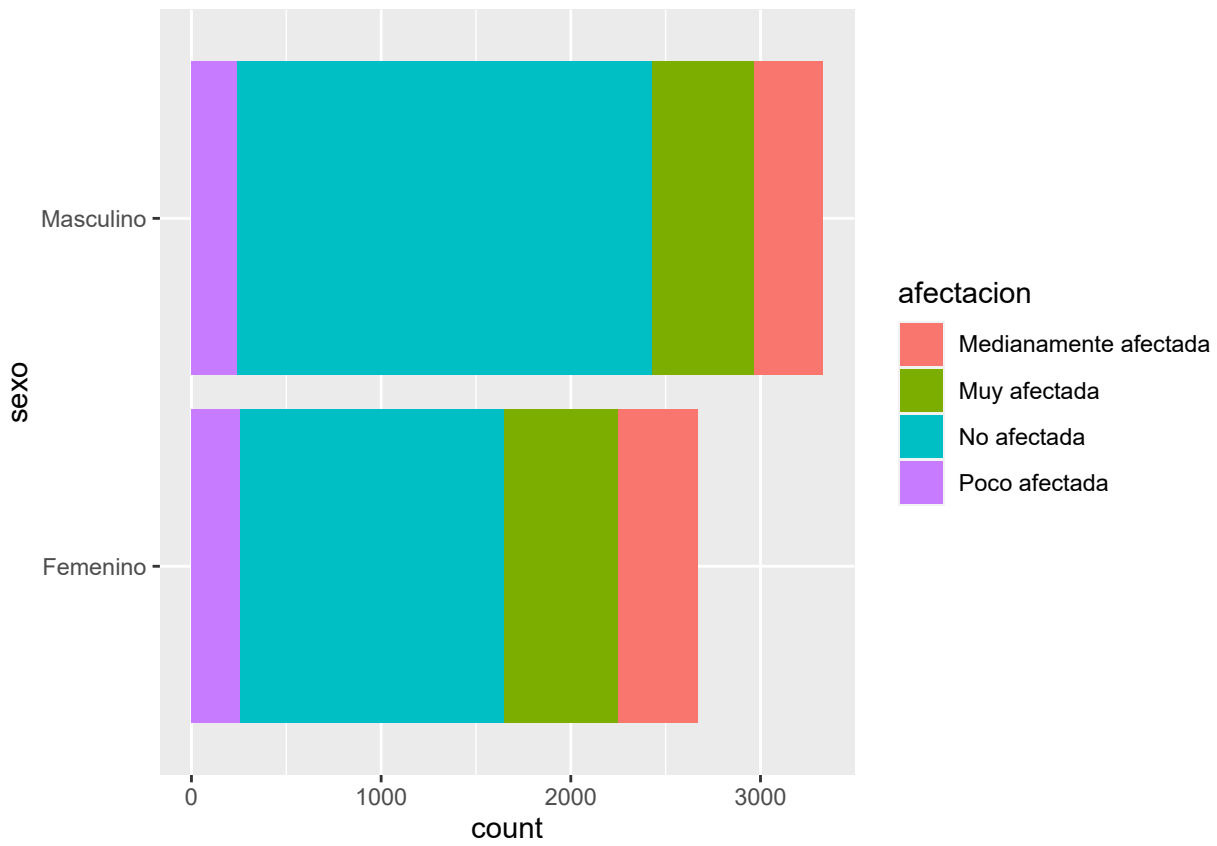
Profundizemos en esto en base a la siguiente página:

[Enlace acá](#)

Podemos voltear el gráfico, y agregar diferenciación en base a otra variable:

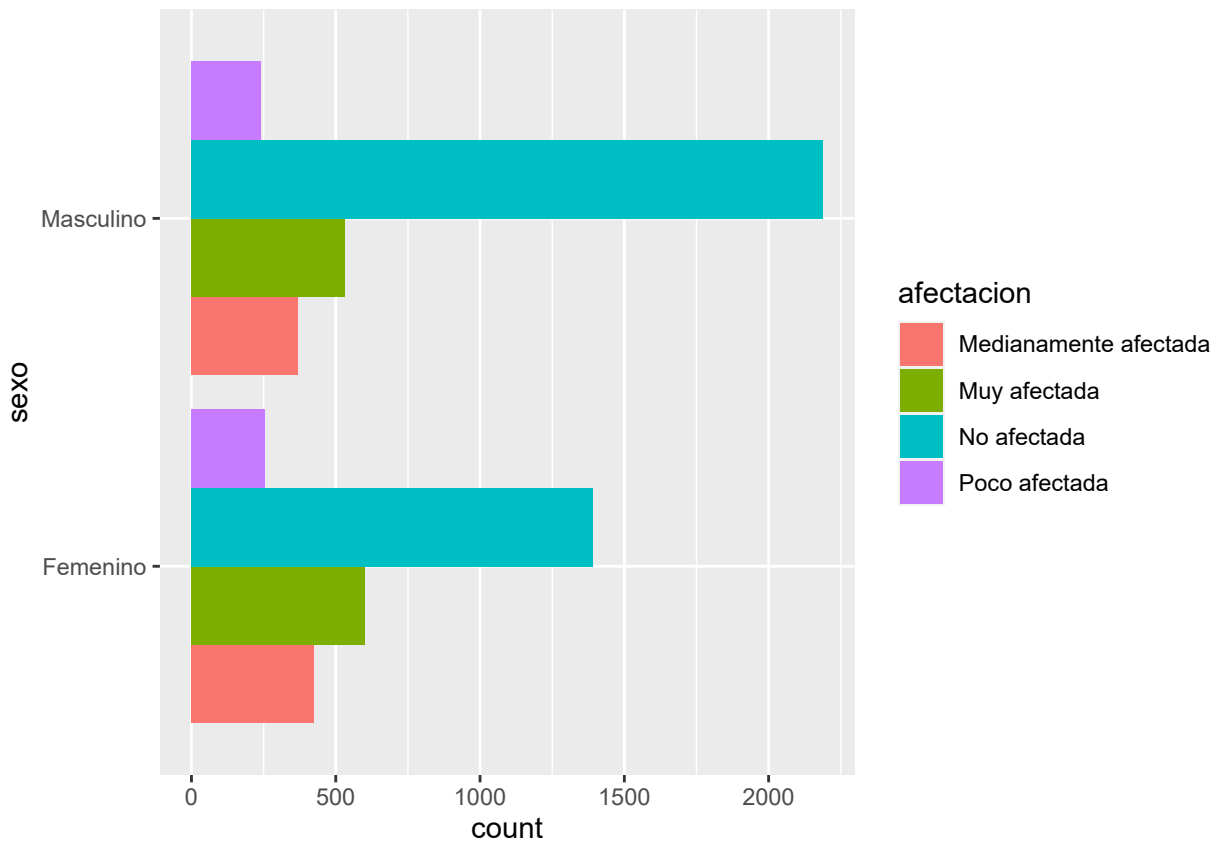
```
p <- ggplot(data) + aes(y = sexo) + geom_bar(aes(fill = afectacion))
```

```
p
```



Podríamos graficar de manera separada cada columna, en base a esta página [Enlace acá:](#)

```
p <- ggplot(data) + aes(y = sexo, fill = afectacion) + geom_bar(position=position_dodge())
p
```



### 3. Fechas Relevantes

Unidad	Evaluación	Ponderación	Fecha
Unidad I	Evaluación diagnóstica		25/03/2021
	Evaluación Individual Participación	(5 %)	05/04/2021
	Evaluación Grupal	(15 %)	27/04/2021 - 04/05/2021
Unidad II	Evaluación Individual - Sumativa I	(30 %)	11/05/2021
	Evaluación Grupal	(15 %)	20/06/2021
	Evaluación Grupal - Sumativa II	(15 %)	04/07/2021
Unidad III	Evaluación Formativa		-
	Evaluación Grupal Sesión I- Sumativa III	(20 %)	08/07/2021
	Evaluación grupal Sesión II- Sumativa III	(20 %)	13/07/2021