

Acuerdos para hacer gráficas: un manual

Manuel Toral

April 3, 2019

Contents

Cosas de cajón	2
Librerías necesarias	2
Datos	2
Diseño muestral	2
Gráficas descriptivas	3
Ejemplo y pasos	4
Gráficas de cruces	9
Ejemplo y pasos	9

Cosas de cajón

Librerías necesarias

```
library(tidyverse) # El set de herramientas que ya conocemos y amamos.  
library(sf) # Para las geografías  
library(foreign) # Para llamar datos de otros formatos.  
library(srvyr) # La librería necesaria para poder tratar encuestas.  
library(Cairo) # Para usar TODAS las fuentes.  
library(hrbrthemes) # Para usar los themes de HRBR.  
extrafont::loadfonts()
```

Datos

```
encuesta <- read.spss("Base Encuesta Nacional Corrupción.sav") %>%  
  as.tibble() # Para convertir la base de datos en un "tibble"
```

Diseño muestral

```
design <- encuesta %>%  
  as_survey_design(ids=1,  
    weight= FE_Final_)
```

Gráficas descriptivas

Ejemplo y pasos

1. Procesar los datos

Como ya expuse en el Manual anterior, el comando `survey_mean()` ofrece las proporciones de la pregunta en cuestión. Para poder obtener los errores estándar y los intervalos de confianza a un nivel convencional del 95% se añaden los argumentos `vartype = c("se","ci")` y `level = 0.95`), **separados siempre uno de otro de una coma (,)**.

En el siguiente *chunk* guardé el procesamiento en el objeto `procesamiento`.

```
design %>%  
  group_by(P9) %>% # Usaré para este ejemplo la pregunta 9.  
  summarise(prop = survey_mean(vartype = c("se","ci"), # Obtener el SE y el CI.  
                                level = 0.95)) %>% # Determinar el nivel 95%.  
  mutate(P9 = recode(P9, #Recodificar los saltos de renglón para que quepan.  
                    "Sólo algunas veces/Casi nunca" =  
                    "Sólo algunas veces/\nCasi nunca", # Se añade "\n"  
                    "La mayoría de las veces" =  
                    "La mayoría\nde las veces")) -> procesamiento  
  
pander::pander(procesamiento)
```

P9	prop	prop_se	prop_low	prop_upp
Sí se justifica	0.1786	0.01085	0.1573	0.1999
No se justifica	0.7868	0.01153	0.7642	0.8094
Ns/Nc	0.03454	0.005059	0.02462	0.04446

2. Hacer el *mapping*

El *mapping* de `ggplot` se compone de una fuente de datos, el objeto de datos `procesamiento`, de una variable x que expresa las respuestas de la pregunta en cuestión, una y que utiliza la variable `prop` que creamos en el paso anteriores, **se rellena**, usando `fill=`, **con las respuestas de la variable** y se construye la etiqueta con `paste0(round(prop*100,1),"%")`.

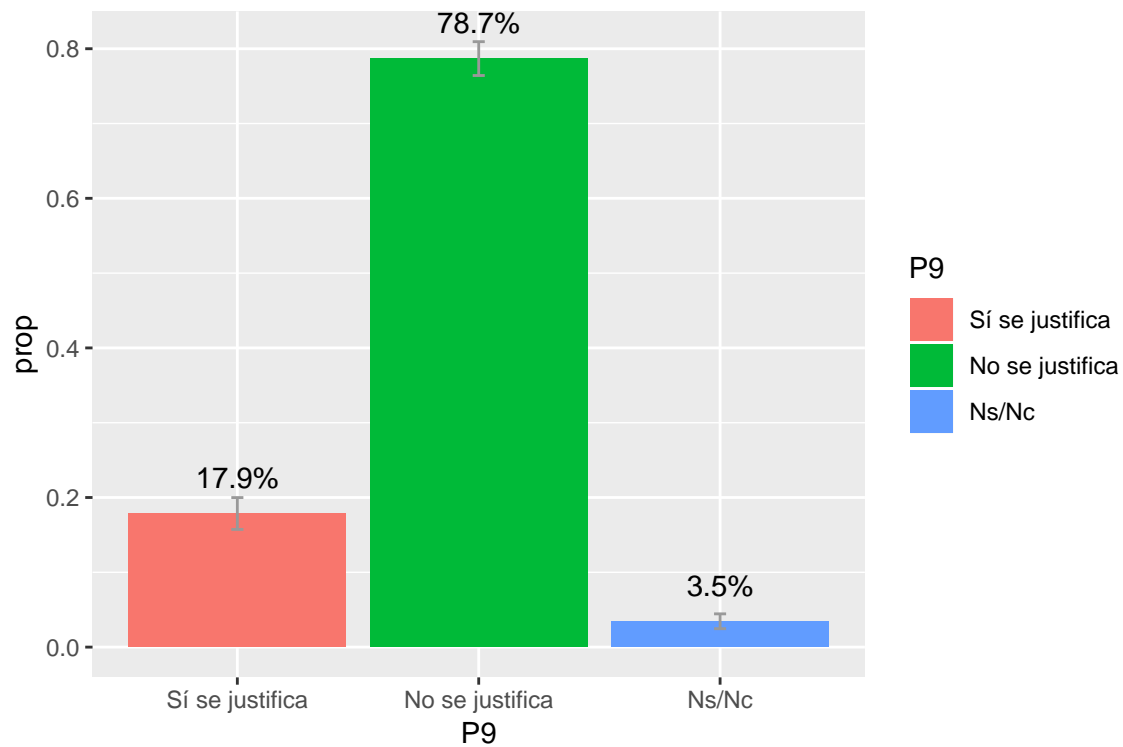
Este último comando a) pega sin espacios con `paste0` el b) redondeo a un decimal de la variable proporción multiplicada por 100 (`round(prop*100,1)`) con c) el caracter `"%"`. Se guarda en el objeto `gp9`.

```
ggplot(procesamiento, aes(x= P9,  
  y= prop,  
  fill= P9,  
  label=paste0(round(prop*100,1),"%"))) -> gp9
```

3. Poner dos geometrías

Vamos a usar tres geometrías, una barra con `geom_col()`, un `geom_text()` para asignar la etiqueta y una barra de error con `geom_errorbar`. Para ésta última se necesitan dos argumentos, naturalmente, la proporción del límite superior `ymin=prop_low` y la del límite inferior `ymax=prop_upp`.

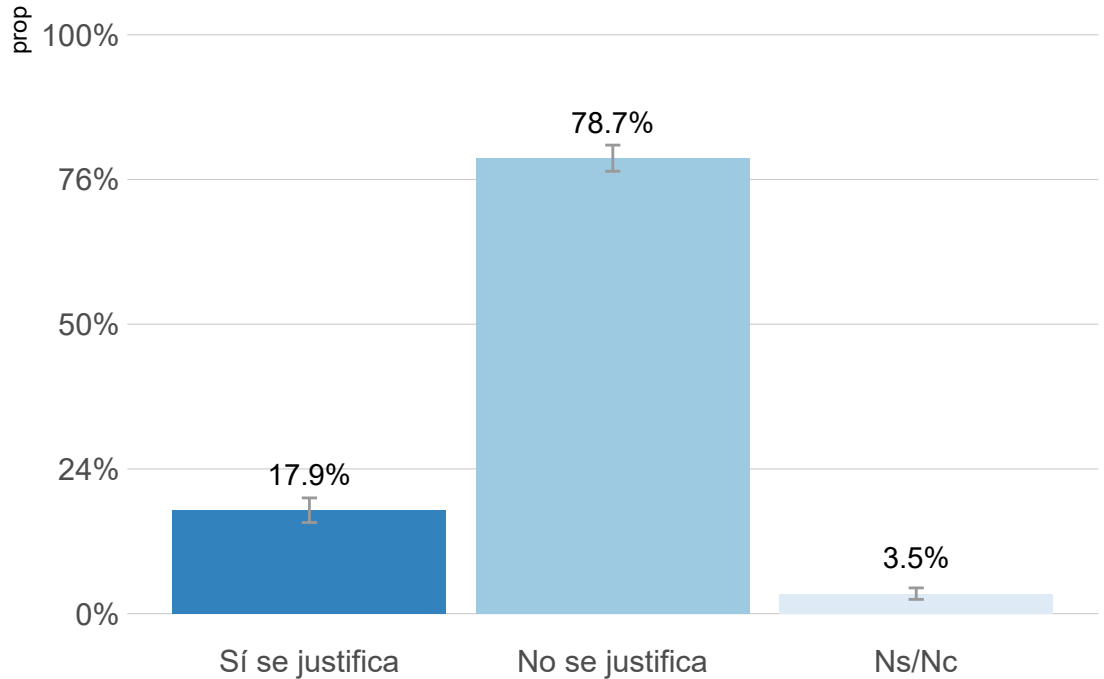
```
gp9 +  
  geom_col() +  
  geom_errorbar(aes(ymin=prop_low,  
                    ymax=prop_upp),  
                width=.05, # La anchura de la barra de error.  
                color="grey60") + # El color de dicha barra.  
  geom_text(vjust=-1.2)-> gp9g # El vjust es para justar horizontalmente.  
gp9g
```



4. Ajustes estéticos

```
gp9g +  
  guides(fill=FALSE) + # Elimina la información redundante de los colores.  
  scale_fill_brewer(palette="Blues", direction = -1) + # Paleta "Blues"  
  scale_y_continuous(labels = # Pone el eje en porcentaje.  
    scales::percent_format(accuracy=2), # 2 dígitos  
    limits=c(0,1)) + # Límites de la gráfica.  
  theme_ipsum(grid = "Y") -> gp9e # Sólo líneas del eje de las Y.
```

gp9e

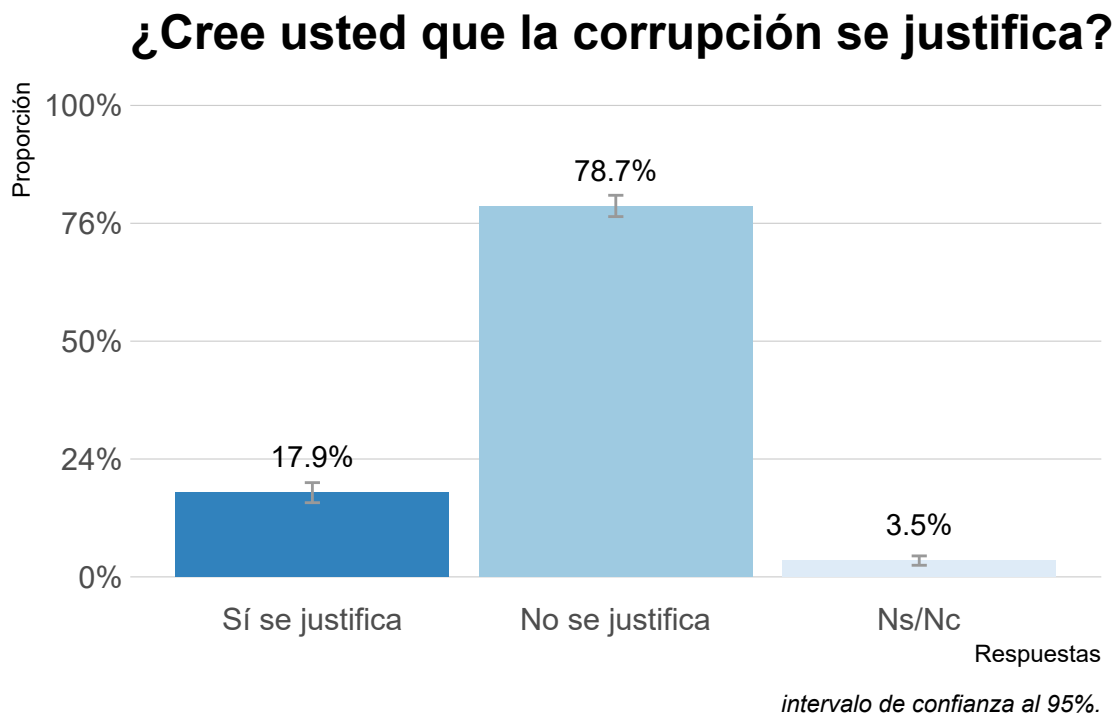


P9

5. Etiquetas

Faltan unas etiquetas que expresen la información contenida en los ejes.

```
gp9e +  
labs(title= "¿Cree usted que la corrupción se justifica?",  
      x="Respuestas",  
      y="Proporción",  
      caption="intervalo de confianza al 95%.")
```



Gráficas de cruces

Ejemplo y pasos

1. Procesar los datos

Ver el paso 1 del proceso anterior. La diferencia fundamental para hacer un cruce es **poner las dos variables en el group_by()**.

En el siguiente *chunk* guardé el procesamiento en el objeto **cruce**.

Cruzaré con género para poder mostrar este ejemplo.

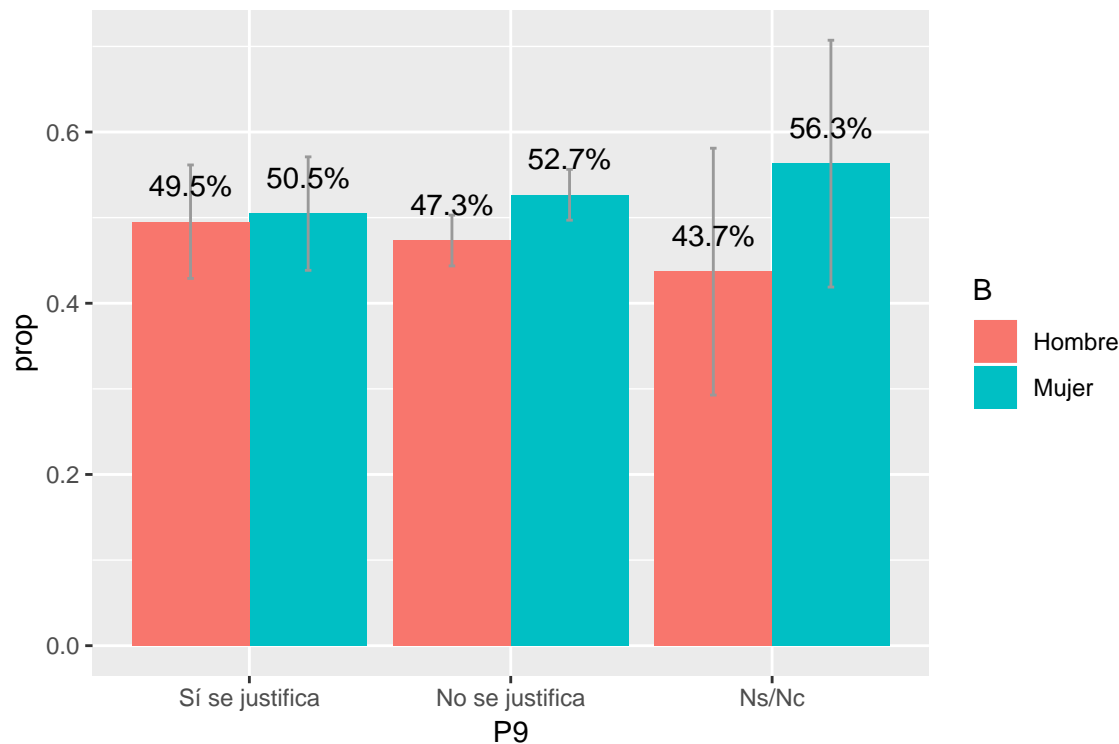
```
design %>%  
  group_by(P9, B) %>% # Usaré para este ejemplo la pregunta 9.  
  summarise(prop = survey_mean(vartype = c("se","ci"), # Obtener el SE y el CI.  
                                level = 0.95)) %>% # Determinar el nivel 95%.  
  mutate(P9 = recode(P9, #Recodificar los saltos de renglón para que quepan.  
                    "Sólo algunas veces/Casi nunca" =  
                    "Sólo algunas veces/\nCasi nunca", # Se añade "\n"  
                    "La mayoría de las veces" =  
                    "La mayoría\nde las veces")) -> cruce  
  
pander::pander(cruce)
```

P9	B	prop	prop_se	prop_low	prop_upp
Sí se justifica	Hombre	0.4953	0.03378	0.429	0.5615
Sí se justifica	Mujer	0.5047	0.03378	0.4385	0.571
No se justifica	Hombre	0.4733	0.01514	0.4436	0.503
No se justifica	Mujer	0.5267	0.01514	0.497	0.5564
Ns/Nc	Hombre	0.4369	0.0735	0.2928	0.5811
Ns/Nc	Mujer	0.5631	0.0735	0.4189	0.7072

2. Procedimiento (abreviado) de graficación

Al ser un proceso similar, haré énfasis en las diferencias para hacer el cruce.

```
ggplot(cruce, aes(x= P9,  
  y= prop,  
  fill= B, # Es la parte fundamental de la tercera pieza de info.  
  label=paste0(round(prop*100,1),"%")) +  
geom_col(position = "dodge") +  
geom_errorbar(aes(ymin=prop_low,  
  ymax=prop_upp),  
  width=.05, # La anchura de la barra de error.  
  color="grey60", # El color de dicha barra.  
  position = position_dodge(0.9)) + # Posicionar barra  
geom_text(vjust=-1.2,  
  position = position_dodge(0.9))-> gp9Bg  
gp9Bg
```

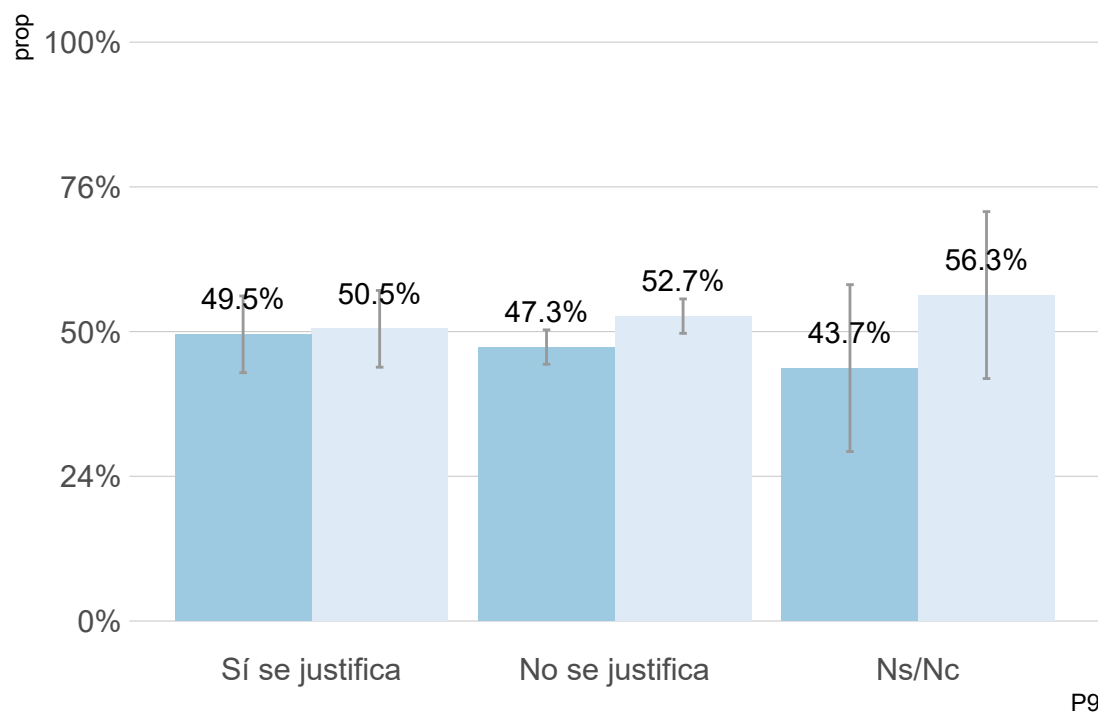


3. Consideraciones estéticas

```
gp9Bg +  
guides(fill=FALSE) + # Elimina la información redundante de los colores.  
scale_fill_brewer(palette="Blues", direction = -1) + # Paleta "Blues"  
scale_y_continuous(labels = # Pone el eje en porcentaje.  
  scales::percent_format(accuracy=2), # 2 dígitos  
  limits=c(0,1))+ # Límites de la gráfica.
```

```
theme_ipsum(grid = "Y") -> gp9Be # Sólo líneas del eje de las Y.
```

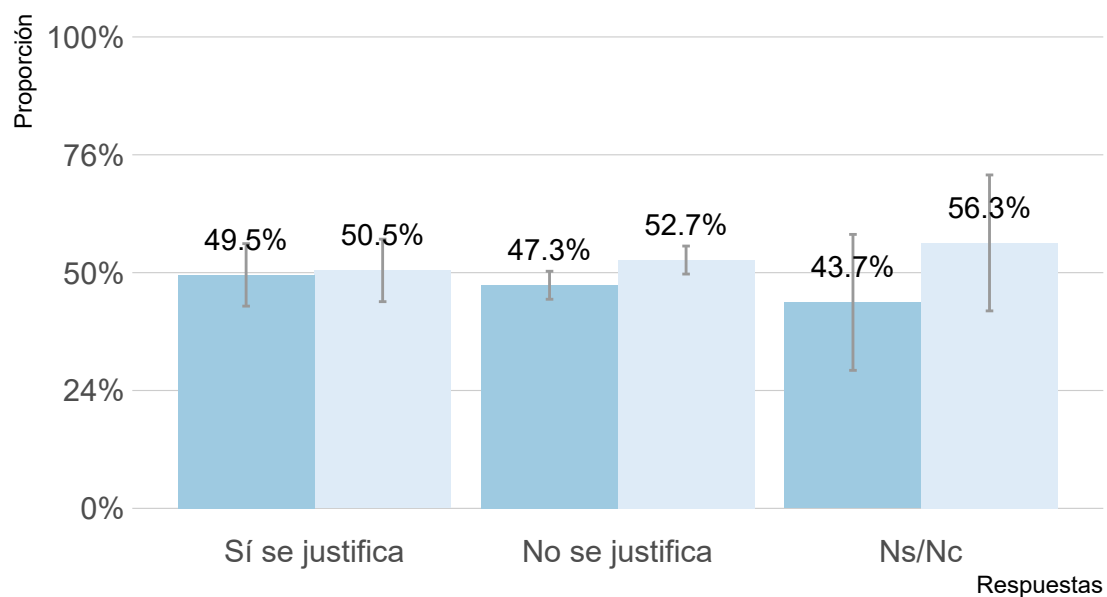
gp9Be



4. Etiquetas

```
gp9Be +
  labs(title= "¿Cree usted que la corrupción se justifica?",
        x="Respuestas",
        y="Proporción",
        caption="intervalo de confianza al 95%.")
```

¿Cree usted que la corrupción se justifica?



intervalo de confianza al 95%.