

Fundamentos de ciencia de datos con R - Módulo 4

Clase 2: Gráficos básicos - barras

CEPAL - Unidad de Estadísticas Sociales

2025-11-04

Introducción

En esta clase trabajaremos gráficos de barras paso a paso con ggplot2: conteos, totales agregados, barras agrupadas, apiladas y proporcionales (100 %), con ordenamientos, etiquetas y opciones de presentación.

Para esta clase utilizaremos la librería tidyverse, que agrupa varias herramientas fundamentales para el análisis de datos en R. Dentro de ella se encuentra ggplot2, el paquete que usaremos para crear nuestros gráficos de barras.

```
library(tidyverse)
```

Introducción

Nota

“La visualización de datos convierte números en argumentos.”
— Adaptado de Tukey

¿Cuándo usar un gráfico de barras?

Usa barras cuando:

- ▶ Comparamas categorías (nominales u ordinales) entre sí.
- ▶ Muestras frecuencias (conteos) o valores agregados (suma, promedio) por categoría.
- ▶ Requieres comparar subgrupos (barras agrupadas o apiladas) dentro de cada categoría.

Evita barras cuando:

- ▶ Quieres mostrar distribuciones de variables continuas (usa histograma/densidad).
- ▶ Es una serie temporal densa y continua (prefiere línea).
- ▶ Quieres mostrar parte–todo dinámico con muchos segmentos (considera treemap o barras ordenadas y limitando categorías).

Carga base de datos ejemplo

```
datos <- readRDS("../Data/base_personas_gasto.rds")
head(datos[,1:8], 5)
```

id_hogar	id_pers	upm	estrato	area	fep	pobreza	ingreso_hh
262	1	1100100006	11001	1	19	3	10783.05
262	2	1100100006	11001	1	19	3	10783.05
265	1	1100100006	11001	1	16	3	21259.72
265	2	1100100006	11001	1	16	3	21259.72
265	3	1100100006	11001	1	16	3	21259.72

Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Paso 1: capa base (solo mapeo estético)

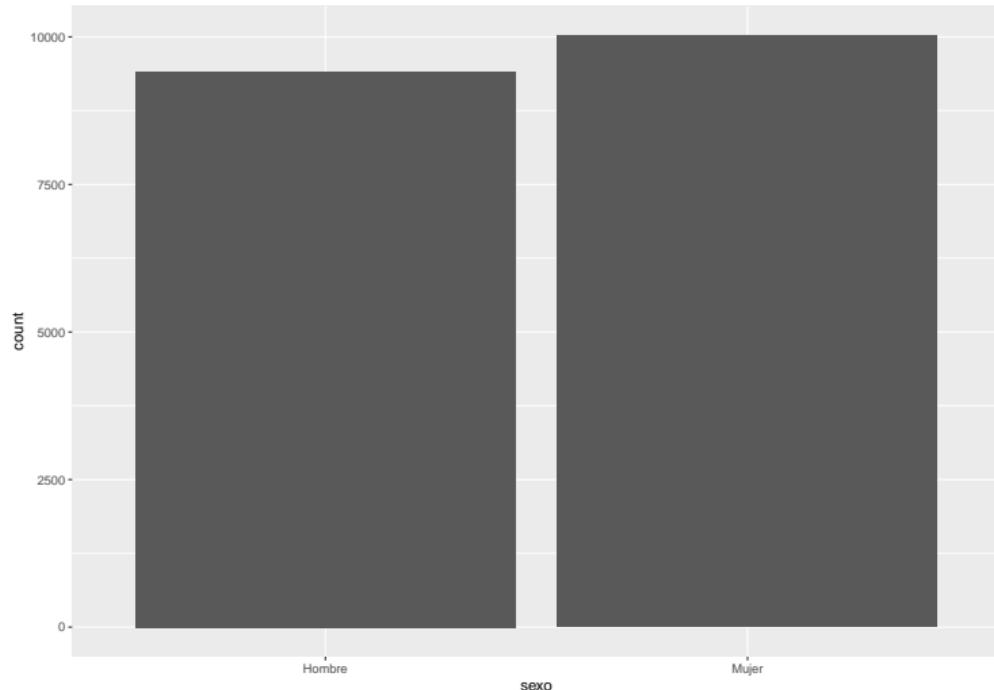
```
ggplot(datos, aes(x = sexo))
```



Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Paso 2: añadir geometría de barras (conteo automático)

```
ggplot(datos, aes(x = sexo) )+ geom_bar()
```



Paso a paso: construyendo un gráfico de barras (conteo por sexo)

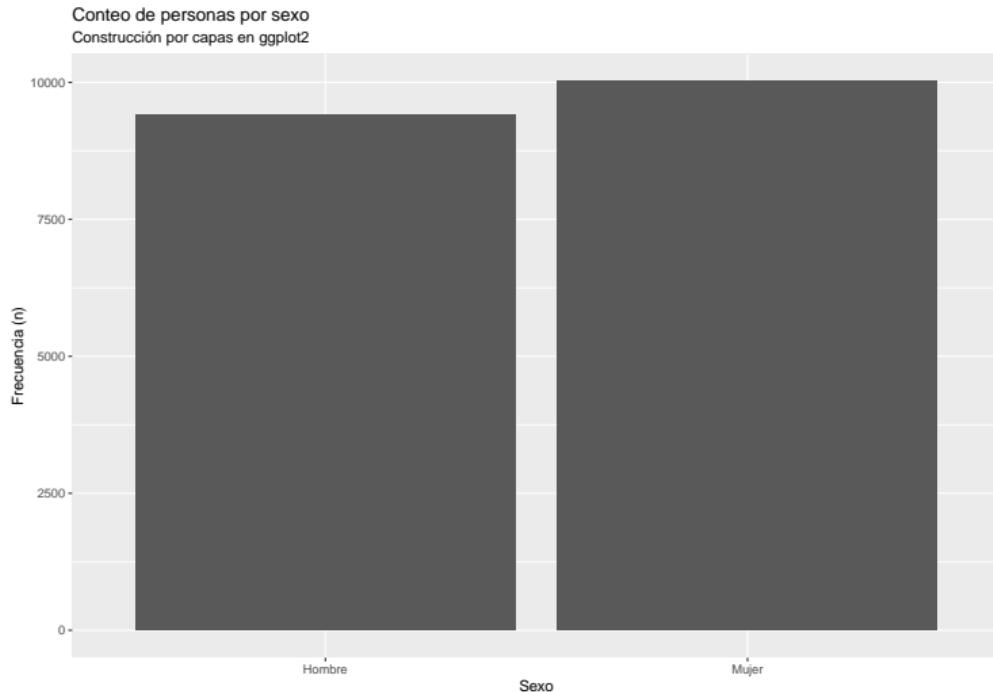
Paso 3: títulos y ejes

```
paso3 <- ggplot(datos, aes(x = sexo) )+ geom_bar() +  
  labs(  
    title = "Conteo de personas por sexo",  
    subtitle = "Construcción por capas en ggplot2",  
    x = "Sexo", y = "Frecuencia (n)"  
  )
```

Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Paso 3: títulos y ejes

paso3



Paso a paso: construyendo un gráfico de barras (conteo por sexo)

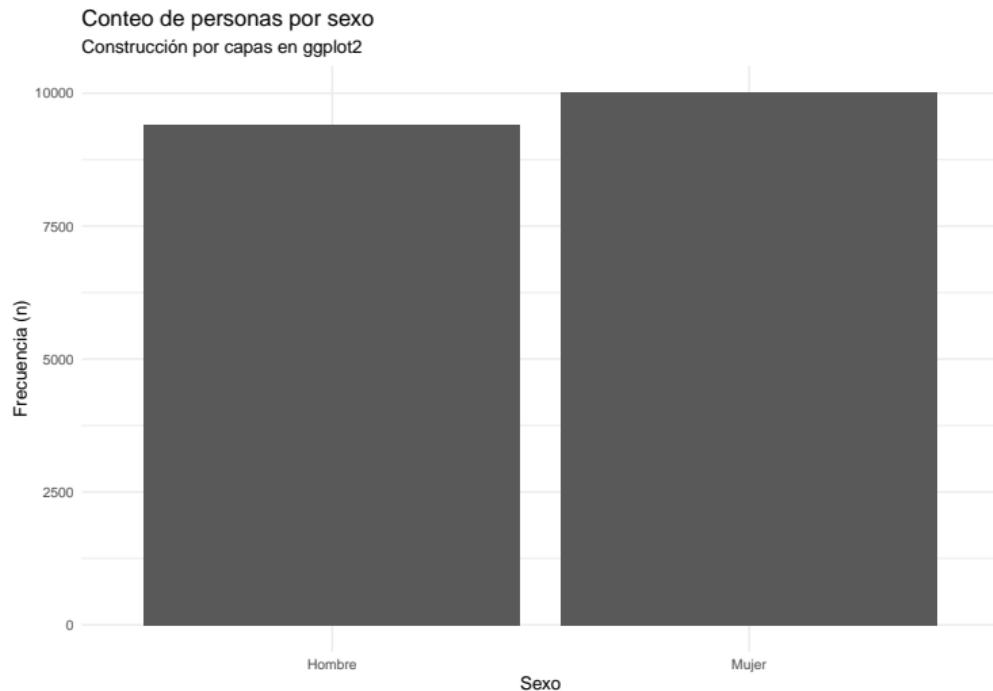
Paso 4: tema y tamaño base

```
paso4 <- ggplot(datos, aes(x = sexo) )+ geom_bar() +  
  labs(  
    title = "Conteo de personas por sexo",  
    subtitle = "Construcción por capas en ggplot2",  
    x = "Sexo", y = "Frecuencia (n)"  
  ) + theme_minimal(base_size = 13)
```

Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Paso 4: tema y tamaño base

paso4



Paso a paso: construyendo un gráfico de barras (conteo por sexo)

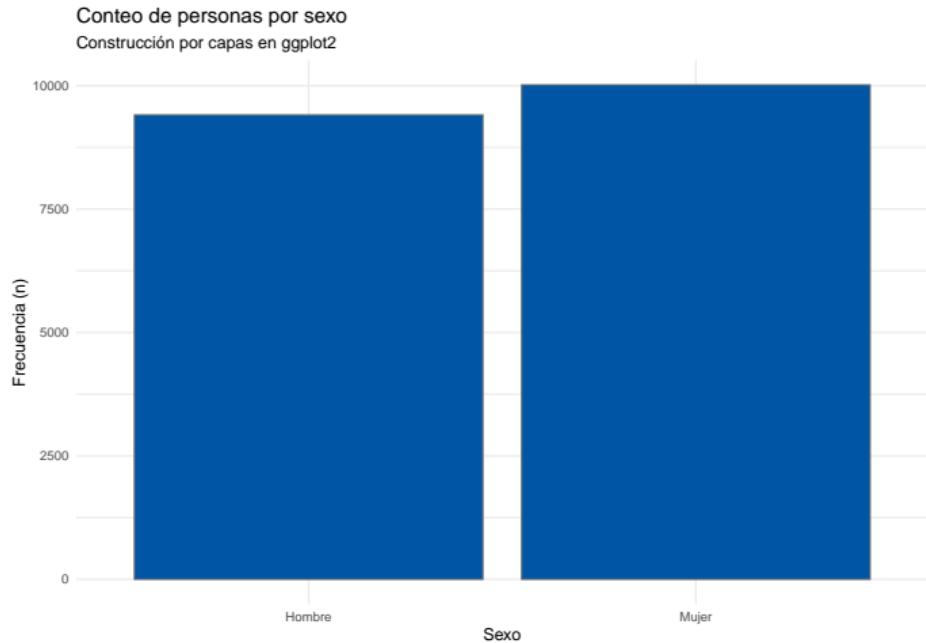
Paso 5: agregar un solo color (relleno fijo)

```
paso5 <- ggplot(datos, aes(x = sexo) )+
  geom_bar(fill = "#0055A4", color = "#6D6E71")+
  labs(title = "Conteo de personas por sexo",
       subtitle = "Construcción por capas en ggplot2",
       x = "Sexo", y = "Frecuencia (n)"
  ) + theme_minimal(base_size = 13)
```

Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Paso 5: a). agregar un solo color (relleno fijo)

paso5



Paso a paso: construyendo un gráfico de barras (conteo por sexo)

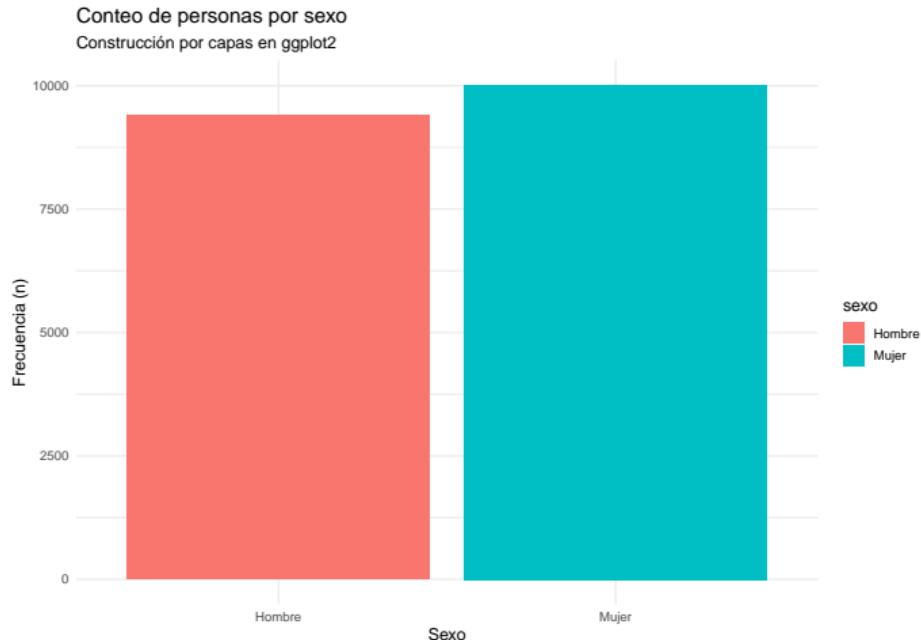
Paso 5: b). un color por categoría (relleno mapeado a la variable)

```
paso5b <- ggplot(datos, aes(x = sexo)) +  
  geom_bar(aes(fill = sexo)) +  
  labs(title = "Conteo de personas por sexo",  
       subtitle = "Construcción por capas en ggplot2",  
       x = "Sexo", y = "Frecuencia (n)")  
  ) + theme_minimal(base_size = 13)
```

Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Paso 5: b). un color por categoría (relleno mapeado a la variable)

paso5b



Paso a paso: construyendo un gráfico de barras (conteo por sexo)

Nota

En gráficos de barras se usa el argumento `fill` (relleno) para definir el color. El argumento `color` solo afecta los bordes de las barras.

Ejemplos adicionales de gráficos de barras

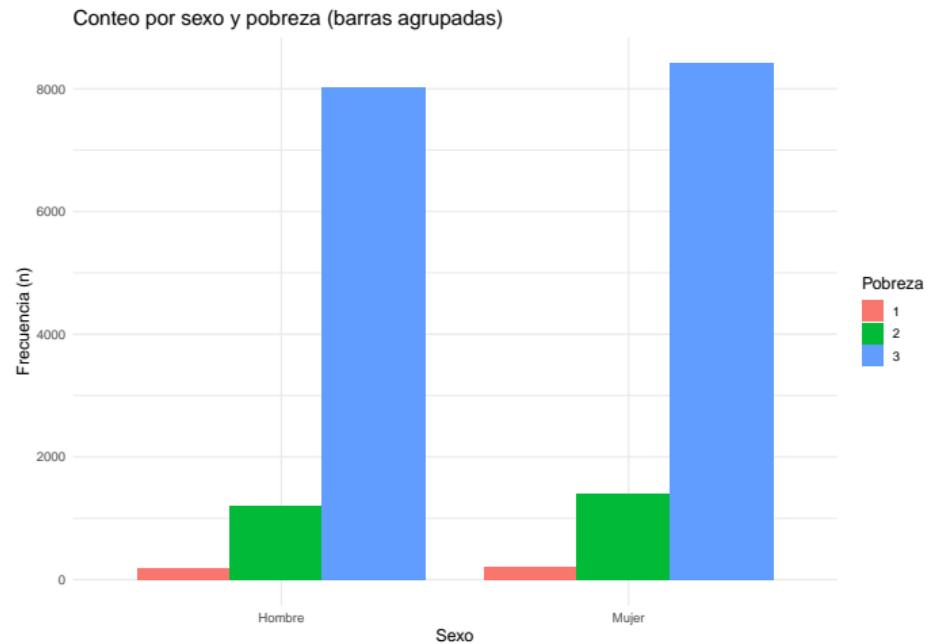
Ejemplo 1: Barras agrupadas (sexo × pobreza)

```
ejem1 <- ggplot(datos, aes(x = sexo, fill = pobreza)) +  
  geom_bar(position = position_dodge(width = 0.8)) +  
  labs(  
    title = "Conteo por sexo y pobreza (barras agrupadas)",  
    x = "Sexo", y = "Frecuencia (n)", fill = "Pobreza"  
  ) +  
  theme_minimal(base_size = 13)
```

Ejemplos adicionales de gráficos de barras

Ejemplo 1: Barras agrupadas (sexo × pobreza)

ejem1



Ejemplos adicionales de gráficos de barras

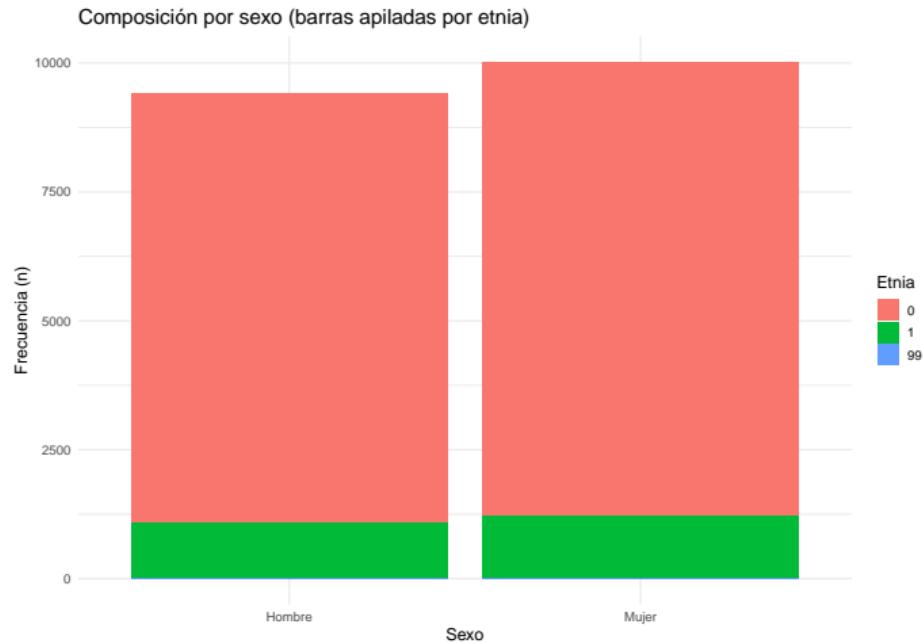
Ejemplo 2: Barras apiladas (composición de etnia por sexo)

```
ejem2 <- ggplot(datos, aes(x = sexo, fill = etnia)) +  
  geom_bar(position = "stack") +  
  labs(  
    title = "Composición por sexo (barras apiladas por etnia)",  
    x = "Sexo", y = "Frecuencia (n)", fill = "Etnia"  
  ) +  
  theme_minimal(base_size = 13)
```

Ejemplos adicionales de gráficos de barras

Ejemplo 2: Barras apiladas (composición de etnia por sexo)

ejem2



Ejemplos adicionales de gráficos de barras

Ejemplo 3: Barras 100 % (proporciones de pobreza dentro de sexo)

Para este ejemplo usaremos la librería scales, que nos permite formatear los ejes y etiquetas como porcentajes (p. ej., con `label_percent()`), ideal cuando usamos `position = "fill"` para mostrar composiciones al 100 %.

```
library(scales)

ejem3 <- ggplot(datos, aes(x = sexo, fill = pobreza)) +
  geom_bar(position = "fill") +
  scale_y_continuous(labels = percent) +
  labs(
    title = "Proporción de pobreza por sexo (barras 100%)",
    x = "Sexo", y = "Porcentaje", fill = "Pobreza"
  ) +
  theme_minimal(base_size = 13)
```

Ejemplos adicionales de gráficos de barras

Ejemplo 3: Barras 100 % (proporciones de pobreza dentro de sexo)

ejem3

