Uso de R para analizar la EPH

Taller de R-Ladies Mendoza

Yaz

Jan

Ana Laura

2023-09-15

# Prerrequisitos

Vamos a importar las librerías con las que trabajar

#### Veamos

library(eph)  
library(tidyverse)  
library(gt)

Ahora, vamos a cargar los datos

# Importamos los datos de la librería EPH  
base\_eph <- eph::toybase\_individual\_2016\_04  
base\_eph <- as\_tibble(base\_eph)  
  
# Veamos 5 registros aleatorios del conjunto de datos  
sample\_n(base\_eph, 5)

# A tibble: 5 × 177  
 CODUSU ANO4 TRIMESTRE NRO\_HOGAR COMPONENTE H15 REGION MAS\_500 AGLOMERADO  
 <fct> <int> <int> <int> <int> <int> <int> <fct> <int>  
1 TQRMNOPR… 2016 4 1 2 1 43 S 4  
2 TQRMNOQY… 2016 4 1 4 0 43 S 5  
3 TQRMNOPV… 2016 4 1 2 1 43 S 5  
4 TQRMNOPX… 2016 4 1 2 1 43 S 34  
5 TQRMNOQX… 2016 4 1 2 1 41 N 15  
# ℹ 168 more variables: PONDERA <int>, CH03 <int>, CH04 <int>, CH05 <fct>,  
# CH06 <int>, CH07 <int>, CH08 <int>, CH09 <int>, CH10 <int>, CH11 <int>,  
# CH12 <int>, CH13 <int>, CH14 <int>, CH15 <int>, CH15\_COD <fct>, CH16 <int>,  
# CH16\_COD <fct>, NIVEL\_ED <int>, ESTADO <int>, CAT\_OCUP <int>,  
# CAT\_INAC <int>, IMPUTA <int>, PP02C1 <int>, PP02C2 <int>, PP02C3 <int>,  
# PP02C4 <int>, PP02C5 <int>, PP02C6 <int>, PP02C7 <int>, PP02C8 <int>,  
# PP02E <int>, PP02H <int>, PP02I <int>, PP03C <int>, PP03D <int>, …

### Algunas demostraciones de lo visto en los talleres

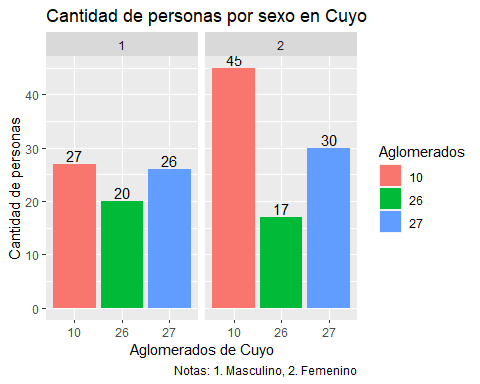
**Manipulación de datos**

# - AGLOMERADO  
# 10 = Gran Mendoza  
# 26 = Gran San Luis  
# 27 = Gran San Juan  
# - Condición de actividad = ESTADO  
# 0 = Entrevista individual no realizada  
# 1 = Ocupado  
# 2 = Desocupado  
# 3 = Inactivo  
# 4 = Menor de 10 años  
cuyo <- base\_eph %>%  
 filter(AGLOMERADO %in% c(10, 26, 27)) %>%  
 count(AGLOMERADO, ESTADO) %>%  
 group\_by(AGLOMERADO) %>%  
 mutate(peso = round(n / sum(n) \* 100, digits = 2))  
  
# Lo convertimos a tabla con gt  
gt(cuyo)

| ESTADO | Casos | peso |
| --- | --- | --- |
| Gran Mendoza | | |
| 1 | 32 | 44.44 |
| 2 | 3 | 4.17 |
| 3 | 25 | 34.72 |
| 4 | 12 | 16.67 |
| 26 | | |
| 1 | 16 | 43.24 |
| 3 | 19 | 51.35 |
| 4 | 2 | 5.41 |
| 27 | | |
| 1 | 19 | 33.93 |
| 2 | 1 | 1.79 |
| 3 | 31 | 55.36 |
| 4 | 5 | 8.93 |

**Visualización de datos**

# Filtro por aglomerados de Cuyo, por sexo  
# CH04 = Sexo: 1 masculino, 2 femenino  
base\_eph %>%  
 filter(AGLOMERADO %in% c(10, 26, 27)) %>%  
 count(AGLOMERADO, CH04) %>%  
 ggplot(aes(x = as.factor(AGLOMERADO),  
 y = n)) +  
 geom\_col(aes(fill = as.factor(AGLOMERADO))) +  
 facet\_wrap(facets = "CH04") +  
 geom\_text(aes(label = n),  
 vjust = -0.25) +  
 labs(title = "Cantidad de personas por sexo en Cuyo",  
 x = "Aglomerados de Cuyo",  
 y = "Cantidad de personas",  
 fill = "Aglomerados",  
 caption = "Notas: 1. Masculino, 2. Femenino")



Fin