

EPH

Psicoestadística. Facultad de Educación. UCC

Alfonso, Matías A.

11 de mayo de 2018

La Encuesta Permanente de Hogares (EPH) es un programa nacional de producción sistemática y permanente de indicadores sociales de propósitos múltiples. Permite conocer características sociodemográficas y socioeconómicas de la población. Actualmente se realiza con una periodicidad de 3 meses y abarca 31 conglomerados urbanos. Descargue la base de datos de la EPH correspondiente al 2do Trimestre de 2017. (<https://www.indec.gob.ar/bases-de-datos.asp>). Abra la base `usu_individual`.

1. Ralice una tabla de distribución de frecuencias y un gráfico de barras para la variable CH12 (“¿Cuál es el nivel más alto que cursa o cursó?”). Indique el nivel de medición de la variable.

```
base %>%  
  group_by(CH12) %>%  
  summarise(freq = n()) %>%  
  mutate(CH12 = factor(CH12,  
                        levels = 1:9,  
                        labels = c("Jardín/Preescolar",  
                                   "Primario",  
                                   "EGB",  
                                   "Secundario",  
                                   "Polimodal",  
                                   "Terciario",  
                                   "Universitario",  
                                   "Posgrado Universitario",  
                                   "Educación Especial")))
```

CH12	freq
NA	3340
Jardín/Preescolar	2119
Primario	16137
EGB	1106
Secundario	22390
Polimodal	670
Terciario	5449
Universitario	7937
Posgrado Universitario	277

CH12	freq
Educación Especial	314
NA	16

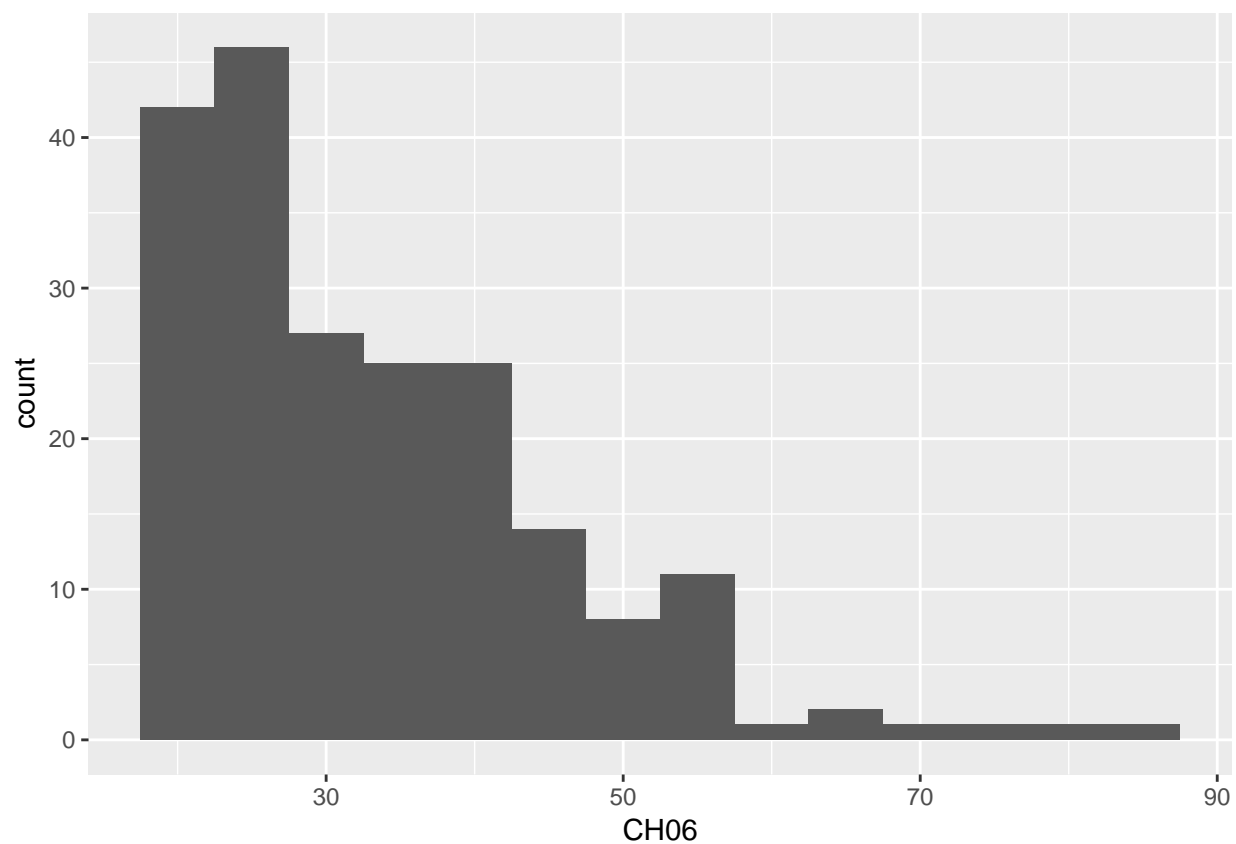
Imagine que está interesado en estudiar la inserción laboral y la capacidad adquisitiva de personas que recibieron educación especial.

2. Seleccione en la base los casos que recibieron Educación Especial y son mayores a 18 años. Construya y guarde una nueva tabla con estos datos. De aquí en adelante trabajará con esa nueva tabla.

```
especial <- base %>%
  filter(CH12 == 9, CH06 >= 18)
```

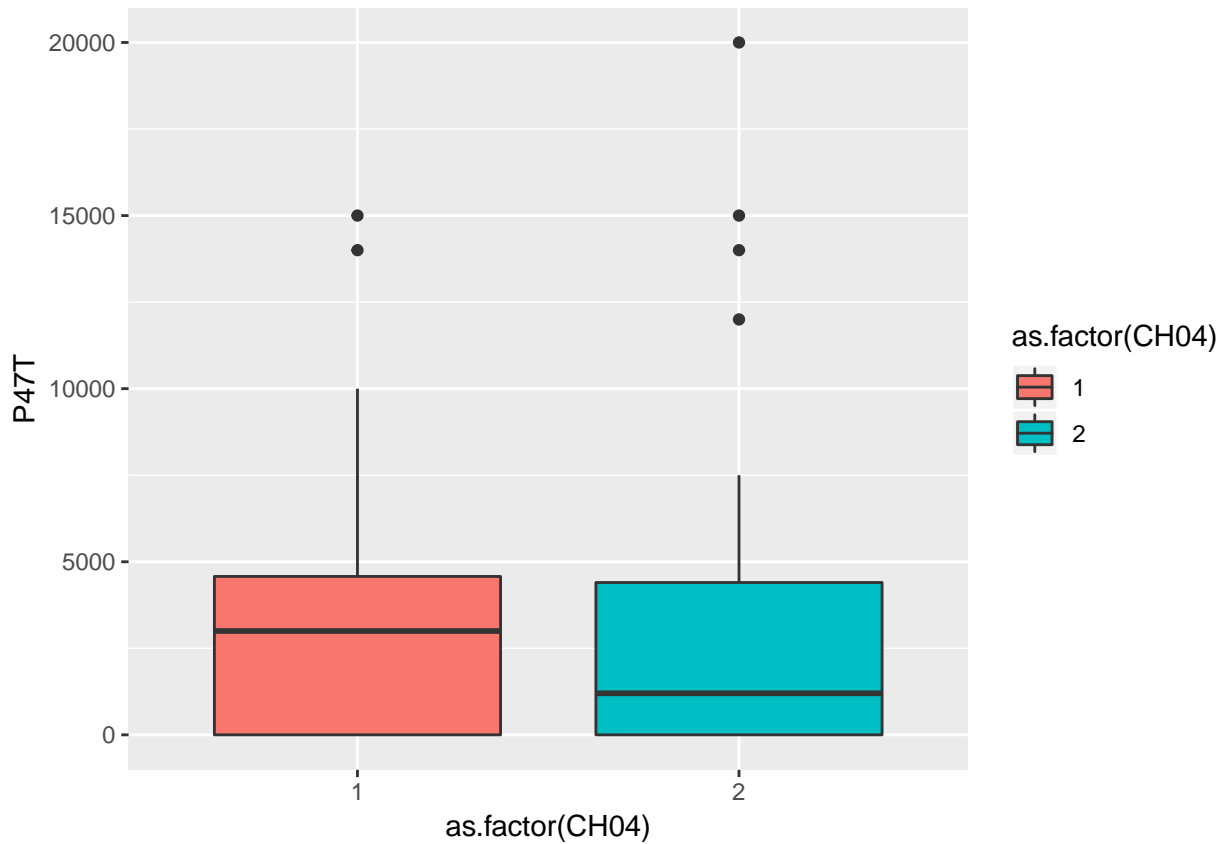
3. Construya un histograma para la Edad (CH06).

```
especial %>%
  ggplot(aes(CH06)) +
  geom_histogram(binwidth = 5)
```



4. Elabore un boxplot para el ingreso total individual (P47T). Interprete los resultados. Indique los cuartiles, la mediana, los valores mínimos y máximos y los resultados extremos.

```
especial %>%
  ggplot(aes(x = as.factor(CH04), y = P47T, fill = as.factor(CH04))) +
  geom_boxplot()
```



5. Realice una tabla de distribución de frecuencias para la variable ESTADO y CAT_INAC. Describa los resultados.

```
especial %>%
  group_by(ESTADO) %>%
  summarise(freq = n()) %>%
  mutate(ESTADO = factor(ESTADO,
                          levels = 1:3,
                          labels = c("Ocupado",
                                      "Desocupado",
                                      "Inactivo")))
```

ESTADO	freq
Ocupado	21
Desocupado	1
Inactivo	183

```

especial %>%
  group_by(CAT_INAC) %>%
  summarise(freq = n()) %>%
  mutate(CAT_INAC = factor(CAT_INAC,
                           levels = 1:7,
                           labels = c("Jubilado/Pensionado",
                                       "Rentista",
                                       "Estudiante",
                                       "Ama de casa",
                                       "Menor de 6 años",
                                       "Discapacitado",
                                       "Otros")))

```

CAT_INAC	freq
NA	22
Jubilado/Pensionado	93
Rentista	1
Estudiante	34
Ama de casa	9
Discapacitado	43
Otros	3