

Ejercicios - Sesión 3

Capacitadores R en DET

14-12-2020

Paso previo

```
## Se cargan los paquetes necesarios para la actividad  
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.3
```

```
## -- Attaching packages -----
```

```
## v ggplot2 3.3.0      v purrr  0.3.3  
## v tibble  3.0.4      v dplyr  1.0.2  
## v tidyr   1.0.0      v stringr 1.4.0  
## v readr   1.3.1      v forcats 0.4.0
```

```
## Warning: package 'ggplot2' was built under R version 3.6.3
```

```
## Warning: package 'tibble' was built under R version 3.6.3
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

```
## -- Conflicts -----
```

```
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()
```

Importar la base **ene-2019-11.csv**.

```
## La base se descarga y se guarda en la carpeta de "inputs" llamada "data"  
ene <- read.csv(file = "data/ene-2019-11.csv")
```

```
## El ejercicio no es del todo correcto si la lectura de la base solo funciona  
# desde la computadora de quien hizo la tarea.
```

Las siguientes tareas realicelas sobre la base ene.

Ejercicio 1

Generar un cuadro de resumen que muestre en las filas la categoría de rama de actividad económica (**b14_rev4cl_caenes**) y en las columnas los grupos ocupacionales (**b1**).

```
# Se agrupa por las dos variables, generando conteo  
cuadro<-ene %>% group_by(b14_rev4cl_caenes,b1) %>% tally()  
cuadro %>% head(10)
```

```
## # A tibble: 10 x 3  
## # Groups:   b14_rev4cl_caenes [2]  
##   b14_rev4cl_caenes    b1      n  
##             <int> <int> <int>
```

```
## 1      1      1      51
## 2      1      2      45
## 3      1      3     141
## 4      1      4      70
## 5      1      5     111
## 6      1      6    1830
## 7      1      7     155
## 8      1      8     300
## 9      1      9    2728
## 10     2      1      28
```

```
# Se aplica pivot_wider() a objeto cuadro
cuadro %>%
  pivot_wider(names_from = b14_rev4cl_caenes,
              names_prefix = "b1_" ,
              values_from = n) %>%

  kableExtra::kable(caption = "Cuadro 1", # Ignorar esta parte, es solo para mejorar visualización en
                    digits = 2, format.args = list( decimal.mark = ",",big.mark = "."),booktabs = T) %>% # Lo mismo
  kableExtra::kable_styling(latex_options = c("scale_down","hold_position")) # Lo mismo
```

Table 1: Cuadro 1

b1	b1_1	b1_2	b1_3	b1_4	b1_5	b1_6	b1_7	b1_8	b1_9	b1_10	b1_11	b1_12	b1_13	b1_14	b1_15	b1_16	b1_17	b1_18	b1_19	b1_20	b1_21	b1_NA
1	51	28	118	11	8	102	171	61	106	31	48	11	37	47	79	123	35	19	12	1	NA	NA
2	45	126	148	22	18	146	142	53	18	141	87	16	506	41	536	1.982	780	77	58	NA	2	NA
3	141	223	318	37	55	273	337	184	89	125	295	123	324	89	570	620	983	106	93	8	1	NA
4	70	48	152	19	16	92	376	223	67	35	88	4	51	103	283	122	108	35	22	NA	1	NA
5	111	73	361	19	14	77	4.468	235	1.191	83	47	106	19	127	373	641	330	117	553	269	1	NA
6	1.830	3	31	2	3	5	42	4	5	NA	NA	2	8	127	49	3	1	12	11	65	NA	NA
7	155	332	1.670	48	34	1.796	600	68	58	41	2	7	39	44	87	26	18	20	245	4	NA	NA
8	300	439	497	26	41	154	251	1.558	37	5	1	1	12	22	124	6	59	5	24	5	NA	NA
9	2.728	185	646	35	106	690	1.213	295	373	21	35	56	28	383	447	359	193	66	192	1.369	1	NA
10	NA	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	201	NA	1	NA	NA	1	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	53.384

Ejercicio 2

Cree las siguientes variables:

- **pet**: que tome valor 1 si edad es mayor o igual a 15, y 0 en otro caso.

```
## Se observa distribución de variable
table(ene$edad)
```

```
##
## 0      1      2      3      4      5      6      7      8      9      10     11     12     13     14     15
## 888 993 1061 1184 1239 1280 1294 1258 1238 1370 1295 1292 1305 1299 1287 1320
## 16   17   18   19   20   21   22   23   24   25   26   27   28   29   30   31
## 1319 1321 1341 1301 1249 1232 1333 1328 1356 1413 1384 1352 1339 1310 1303 1208
## 32   33   34   35   36   37   38   39   40   41   42   43   44   45   46   47
## 1119 1117 1049 1021 1088 1165 1175 1122 1132 1083 1112 1244 1278 1235 1260 1311
## 48   49   50   51   52   53   54   55   56   57   58   59   60   61   62   63
## 1244 1273 1343 1290 1299 1394 1398 1408 1433 1324 1285 1276 1292 1180 1167 1215
## 64   65   66   67   68   69   70   71   72   73   74   75   76   77   78   79
## 1124 1133 1023 950  905  896  847  822  751  805  770  649  635  547  544  500
## 80   81   82   83   84   85   86   87   88   89   90   91   92   93   94   95
## 424  362  336  297  275  247  229  220  188  168  133  115  67   63   51   37
```

```
##    96    97    98    99   100   101   102   103   104
##    43    17    17     7     9     5     1     1     3
```

```
## Se recodifica con mutate() e if_else()
ene %>% mutate(pet=if_else(edad>=15,1,0)) %>% select(pet) %>% sum()
```

```
## [1] 77957
```

```
ene<-ene %>% mutate(pet=if_else(edad>=15,1,0))
table(ene$pet)
```

```
##
##      0      1
## 18283 77957
```

- **ocu**: que tome el valor 1 si la variable cae_especifico se encuentra en el rango (extremos incluidos) entre 1 y 7, y que tome el valor 0 en cualquier otro caso.

```
table(ene$cae_especifico,useNA = "ifany")
```

```
##
##      0      1      2      3      4      5      6      7      8      9     10     11     12
## 18283 38524  421  1114    81  1724   782  210  2584  397  154  243  1113
##    13    14    15    16    17    18    19    20    21    22    23    24    25
##  6504   982  7346   256  3611   150  3708  212  3670  205  523  1637  497
##     26    27    28
##    512   620   177
```

```
## Para varias se utilizará case_when()
ene<-ene %>% mutate(ocu=case_when(cae_especifico>=1 & cae_especifico<=7 ~ 1,
                                  TRUE~0))
```

```
table(ene$ocu,useNA = "ifany")
```

```
##
##      0      1
## 53384 42856
```

Ejercicio 3

Genere un cuadro de resumen agrupado por región que muestre el total de la variable **pet** y el total de la variable **ocu**.

```
cuadro2<-ene %>% group_by(pet,ocu) %>% tally() %>% as.data.frame()## Solo correr esta línea. El resto e
```

```
a<-c("Menores de 15 y no ocupados","15 o más años y no ocupados","15 o más años y ocupados")
rownames(cuadro2)<-a
```

```
cuadro2 %>%
  kableExtra::kable(caption = "Cuadro 2",

  digits = 2, format.args = list( decimal.mark = ",",big.mark = "."),booktabs = T) %>%
  kableExtra::kable_styling(latex_options = c("hold_position"))
```

Table 2: Cuadro 2

	pet	ocu	n
Menores de 15 y no ocupados	0	0	18.283
15 o más años y no ocupados	1	0	35.101
15 o más años y ocupados	1	1	42.856

Ejercicio 4

Sobre el cuadro de resumen del punto anterior, cree la variable **to**, que tomará el valor de: total de ocu / total de pet * 100

```
cuadro2 %>%
  pivot_wider(names_from = c(pet,ocu),names_prefix = "x",values_from = n) %>%
  mutate(to=x1_1/(x1_0+x1_1))
```

```
## # A tibble: 1 x 4
##   x0_0 x1_0 x1_1   to
##   <int> <int> <int> <dbl>
## 1 18283 35101 42856 0.550
```

(De esta forma se obtiene la tasa de ocupación, pero no será equivalente a la oficial pues no se están considerando los factores de expansión.).

Sobre el envío

Enviar a más tardar el día jueves 17 de diciembre a las 17:00. Enviar ejercicios como archivo .R (script) con nombre y apellido de quien hizo la tarea a los correos nicolas.ratto@ine.cl y gonzalo.franetovic@ine.cl.