

Ejercicios - Sesión 3

Capacitadores R en DET

14-12-2020

Pasos previos

Paquetes

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

Base de datos

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  
ej1 <- ene %>%  
  group_by(b14_rev4cl_caenes, b1) %>%  
  count()  
  
## Se previsualiza el objeto  
ej1 %>%  
  head(10)
```

```
## # A tibble: 10 x 3  
## # Groups:   b14_rev4cl_caenes, b1 [10]  
##   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
ej1 <- ej1 %>%
  pivot_wider(names_from = b14_rev4cl_caenes,
              names_prefix = "b1_" ,
              values_from = n)
```

```
## Se despliega el objeto
ej1
```

```
## # A tibble: 11 x 23
## # Groups:   b1 [11]
##      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
##      <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int> <int>
## 1      1     51     28    118     11      8    102    171     61    106     31     48     11
## 2      2     45    126    148     22     18    146    142     53     18    141     87     16
## 3      3    141    223    318     37     55    273    337    184     89    125    295    123
## 4      4     70     48    152     19     16     92    376    223     67     35     88      4
## 5      5    111     73    361     19     14     77   4468    235   1191     83     47    106
## 6      6   1830      3     31      2      3      5     42      4      5     NA     NA      2
## 7      7    155    332   1670     48     34   1796    600     68     58     41      2      7
## 8      8     300    439    497     26     41    154    251   1558     37      5      1      1
## 9      9   2728    185    646     35    106    690   1213    295    373     21     35     56
## 10     10     NA     NA      1     NA     NA     NA     NA     NA     NA     NA     NA     NA
## 11     NA     NA     NA     NA     NA     NA     NA     NA     NA     NA     NA     NA     NA
## # ... with 10 more variables: b1_13 <int>, b1_14 <int>, b1_15 <int>,
## #   b1_16 <int>, b1_17 <int>, b1_18 <int>, b1_19 <int>, b1_20 <int>,
## #   b1_21 <int>, b1_NA <int>
```

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, useNA = "ifany")
```

```
##
##  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()
```

```
ej2 <- ene %>%
  mutate(pet=if_else(edad>=15,1,0))
```

```
## Se despliega la tabla
```

```
table(ej2$pet, useNA = "ifany")
```

```
##
## 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.

```
## Se observa distribución de variable
```

```
table(ej2$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
```

```
## Se recodifica con mutate() y case_when(), para variar
```

```
ej2 <- ej2 %>%
  mutate(ocu=case_when(cae_especifico>=1 & cae_especifico<=7 ~ 1,
    TRUE~0))
```

```
## Se despliega la tabla
```

```
table(ej2$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**.

```
## Se agrupa según variable
ej3 <- ej2 %>%
  group_by(region) %>%
  summarise(pet_total = sum(pet),
            ocu_total = sum(ocu)) %>%
  select(region, pet_total, ocu_total)
```

```
## Se despliega el objeto
ej3
```

```
## # A tibble: 16 x 3
##   region pet_total ocu_total
##   <int>     <dbl>     <dbl>
## 1     1         2383         1347
## 2     2         2585         1481
## 3     3         2195         1250
## 4     4         4572         2547
## 5     5        10266         5420
## 6     6         4638         2550
## 7     7         5345         2944
## 8     8         8087         3928
## 9     9         4139         2193
## 10    10         6076         3277
## 11    11         2164         1441
## 12    12         1767         1101
## 13    13        15937         9060
## 14    14         2838         1600
## 15    15         2665         1516
## 16    16         2300         1201
```

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

```
## Se crea la nueva variable
ej4 = ej3 %>%
  mutate(to = ocu_total/pet_total*100)
```

```
## Se despliega el objeto
ej4
```

```
## # A tibble: 16 x 4
##   region pet_total ocu_total   to
##   <int>     <dbl>     <dbl> <dbl>
## 1     1         2383         1347  56.5
```

##	2	2	2585	1481	57.3
##	3	3	2195	1250	56.9
##	4	4	4572	2547	55.7
##	5	5	10266	5420	52.8
##	6	6	4638	2550	55.0
##	7	7	5345	2944	55.1
##	8	8	8087	3928	48.6
##	9	9	4139	2193	53.0
##	10	10	6076	3277	53.9
##	11	11	2164	1441	66.6
##	12	12	1767	1101	62.3
##	13	13	15937	9060	56.8
##	14	14	2838	1600	56.4
##	15	15	2665	1516	56.9
##	16	16	2300	1201	52.2

(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. Los ejercicios se deben enviar 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`.