

Tarea 1- Fundamentos de programación

Jairo Nicolas Cabrera

8/29/2021

#Solución.

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5    v purrr  0.3.4
## v tibble  3.1.2    v dplyr  1.0.7
## v tidyr   1.1.3    v stringr 1.4.0
## v readr   1.4.0    v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

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

```
library(magrittr)
```

```
##
```

```
## Attaching package: 'magrittr'
```

```
## The following object is masked from 'package:purrr':
```

```
##
```

```
##      set_names
```

```
## The following object is masked from 'package:tidyr':
```

```
##
```

```
##      extract
```

```
library(dplyr)
```

```
cg=read.csv2("C:/Users/jnico/OneDrive/Documents/FD/Talleres/Datos/Caracteristicas Generales.csv")
```

```
ft=read.csv2("C:/Users/jnico/OneDrive/Documents/FD/Talleres/Datos/Fuerza de Trabajo.csv")
```

##Punto 1.

a. Una el módulo Área de características generales (personas) y el módulo Área fuerza de trabajo ¿Cuáles variables deberían usar como llave?

```
dato=left_join(cg,ft,by=c("i..DIRECTORIO","SECUENCIA_P","ORDEN"))
```

Las variables son DIRECTORIO, SECCION_P y ORDEN.

b. Verifique que el tamaño de la nueva base de datos tenga la misma cantidad de registros que la de características de las personas.

```
nrow(dato)
```

```
## [1] 27730
```

```
nrow(cg)
```

```
## [1] 27730
```

c. Usando la librería dplyr, asigne el valor 1 a aquellos que en la pregunta P6240 respondieron 1 y cero a las demás alternativas respondidas. ¿Qué nombre le pondrá al objeto creado?

```
dato %<>% mutate(P6240a=ifelse(P6240==1,1,0))
```

El objeto sera concido como P6240a.

d. A partir del objeto que acaba de crear, cree otro objeto que guarde, agrupado por hogar(DIRECTORIO, SECUENCIA_P) la suma de ocupados en dicho hogar. Asegúrese de sólo dejar los campos llave, y la variable que acaba de crear)

```
oc = dato %>%group_by(i..DIRECTORIO,SECUENCIA_P,P6240a) %>% summarise(Ocupados=n()) %>% filter(P6240a==1)
```

```
## `summarise()` has grouped output by 'i..DIRECTORIO', 'SECUENCIA_P'. You can override using the `.groups` argument.
```

e. Asegúrese de que la variable que almacena el número de ocupados por hogar sea un factor.

```
oc1 =factor(oc$Ocupados)
class(oc1)
```

```
## [1] "factor"
```

f. Presente una tabla que indique cuántos hogares tienen 0 ó 1 ó 2 ó . . . ocupados en dicho hogar. Tenga presente que si cada hogar se repite tantas veces como miembros tiene, debe eliminar los registros repetidos de dicho vector (tal vez con el comando unique())

```
t1 =table(oc$Ocupados)
t1
```

```
##
##      1      2      3      4      5      6      7
## 4046 2013  427   88   16    3    1
```

g. Repita el paso ya ejecutado de asignar el valor 1 a aquellos que en la pregunta P6240 respondieron 1 y cero a las demás alternativas respondidas.

```
dato %<>% mutate(P6240b=ifelse(P6240==1,1,0))
```

h. A partir del objeto que acaba de crear, cree otro objeto que guarde, agrupado por hogar (DIRECTORIO, SECUENCIA_P) la suma de ocupados en dicho hogar. Asegúrese de sólo dejar los campos llave, la variable que acaba de crear y el factor de expansión (fex_c_2011.x)

```
oc2= dato %>% group_by(i..DIRECTORIO,SECUENCIA_P,fex_c_2011.x,P6240b) %>% summarise(ocupados2=n()) %>% :
```

```
## `summarise()` has grouped output by 'i..DIRECTORIO', 'SECUENCIA_P', 'fex_c_2011.x'. You can override
```

i. Asegúrese de que la variable que almacena el número de ocupados por hogar sea un factor.

```
oc3= factor(oc2$P6240b)
class(oc3)
```

```
## [1] "factor"
```

j. Utilice la función wtd.table() donde el primer parámetro es la variable que alberga el número de individuos ocupados por hogar y la segunda variable es el factor de expansión. El tipo de salida que se espera es 'table'

```
library(Hmisc)
```

```
## Warning: package 'Hmisc' was built under R version 4.1.1
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
##
```

```
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:dplyr':
```

```
##
```

```
##      src, summarize
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      format.pval, units
```

```
t2 =wtd.table(oc2$ocupados2,oc2$fex_c_2011.x,"table")
```

```
t2
```

```
##           1           2           3           4           5           6
## 3415802.3567 1745133.9059 335728.8170 69467.2335 13040.5149 2042.2452
##           7
##      709.7465
```

k. Calcule la frecuencia relativa y escriba su interpretación del resultado.

```
prop.table(t2,NULL)
```

```
##           1           2           3           4           5           6
## 0.6119398715 0.3126401666 0.0601457074 0.0124450321 0.0023362040 0.0003658676
##           7
## 0.0001271509
```

##Punto 2.

```
eq=read.csv2("C:/Users/jnico/Downloads/Equivalencias_ciiu.csv")
o= read.csv("C:/Users/jnico/OneDrive/Documents/FD/Talleres/Datos/Ocupado.csv",sep = ";")
```

a. Una el módulo Área de características generales (personas) y el módulo Ocupados, encargándose que quede del tamaño del módulo de ocupados.

```
dato2 =left_join(o,cg, by= c("i..DIRECTORIO","SECUENCIA_P","ORDEN"))
```

b. A continuación, una el archivo de equivalencias ciiu con el objeto de determinar la Sección (variable RAMA2D_R4) en la que trabaja cada individuo ocupado.

```
names(eq)[3] = "RAMA2D_R4"
dato3 =full_join(dato2,eq,by="RAMA2D_R4")
```

c. Presente una tabla con el número de encuestados agrupado por rama de actividad económica Pero esta solución no hace uso de los factores de expansión de la encuesta.

```
t3= table(dato3$Sección)
```

d. Presente una tabla donde se ha sumado el factor de expansión (fex_c_2011.x), agrupado por Sección (Área de la economía).

```
h =dato3 %>% group_by(fex_c_2011.x) %>% select(RAMA2D_R4,Sección,fex_c_2011.x)
t4= table(h$Sección,h$fex_c_2011.x)
```

e. Presente una tabla de la frecuencia relativa de los ocupados por área de la economía y escriba su interpretación del resultado.

```
t5= prop.table(t4)
```