

Practica 3

Operadores lógicos y comparativos.

Paso 1

Vamos a usar una base de datos en formato “.csv” que está en este sitio:

https://ytliu0.github.io/Stat390EF-R-Independent-Study-archive/RMarkdownExercises/confounder_revisited.html

Descargarla en el directorio de trabajo del proyecto.

Usa el Snippets que creaste “my_readCsv” para cargar archivos en “.csv”. Asigna un nombre al nuevo objeto “survey” que será del tipo “dataframe”.

Paso 2

Identifica los operadores lógicos y comparativos.

Operador	Nombre	Ejemplo
==	Igual	x == y
!=	No igual	x != y
>	Mayor que	x > y
<	Menor que	x < y
>=	Mayor que o igual a	x >= y
<=	Menor que o igual a	x <= y

Operador	Descripción
&	Operador lógico "y" por elemento por elemento.
&&	Logical AND operator - Returns TRUE if both statements are TRUE
	Operador lógico "o" por elemento por elemento.
	Logical OR operator. It returns TRUE if one of the statement is TRUE.
!	Negación lógica, Falso si el enunciado es Verdadero

Ejemplo de elemento por elemento a evaluar el enunciado (correrlo):

```
x <- 1:3
x > 1 & x < 3
[1] FALSE TRUE FALSE
x > 1 && x < 3
[1] FALSE
```

Paso 3 Elabora el siguiente ejercicio

```
1 ## Sección
2 ## Fecha: Wed Aug 7 20:02:40 2024
3 #####
4 survey<- read.csv("Stat100_2017spring_survey01M2.csv",
5                 header=T,
6                 na.strings= c('NA', " ", "", "NO TIENE", -99),
7                 quote="\\"",
8                 dec=".",
9                 stringsAsFactors = FALSE)
10 ### Dimensiones de la base de datos
11 dim(survey)
12 ### nombre de las variables
13 names(survey)
14 ### Exploramos una variable
15 survey$gender
16 class(survey$gender)
17 # [1] "character"
18 survey$gender = as.factor(survey$gender)
19 class(survey$gender)
20 levels(survey$gender)
21
22 ### Promedio por género "shoeSize"
23 ZapatoM = mean(survey[survey$gender == 'female', 'shoeSize'])
24 ZapatoH = mean(survey[survey$gender == 'male', 'shoeSize'])
25
26 ZapatoM == ZapatoH
27 ZapatoM >= ZapatoH
28 ZapatoM <= ZapatoH
29 ZapatoM != ZapatoH
30 ZapatoM > ZapatoH
31 ZapatoM < ZapatoH
32
33 # ¿Qué puedes decir de la diferencia?
34 # ¿Cuál es el promedio de hombre y el de mujer?
35 # ¿Son comparables la medición de zapatos de mujer y la de hombre en escala de USA?
36 # Repite el ejercicio de arriba para la variable "shoeNums: How many pairs of shoes do you own?"
```

Paso 4 Creamos una variable

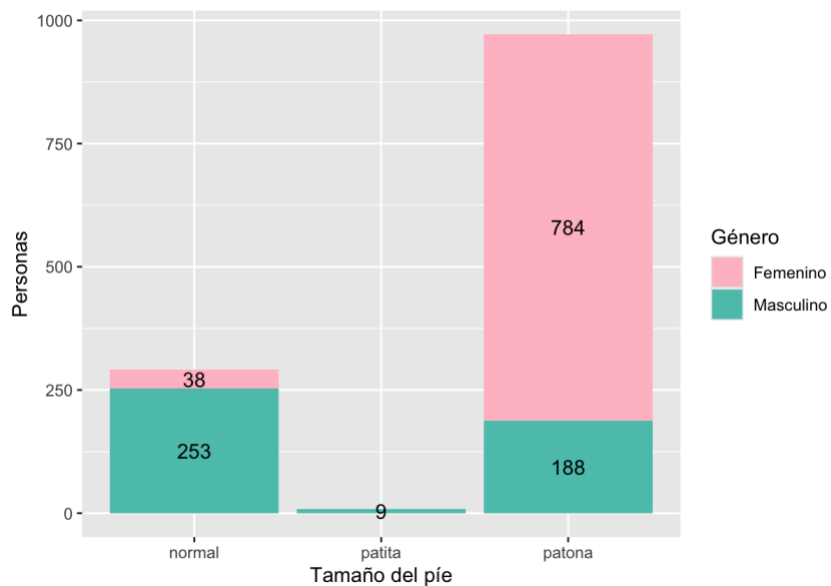
```
40 ## ¿Como crear una nueva variable?
41
42 survey$yeti = ifelse(survey$gender == "male", "pendiente", ifelse(survey$shoeSize >= 6,
43   "patona", ifelse(((survey$shoeSize < 6) & (survey$shoeSize >= 4)), "normal", "patita" )))
44
45 ## Checa si los valores son los correctos
46 survey[1,]
47 survey[2,]
48 survey[3,]
49 survey[4,]
50
51 # [1] "patona"    "pendiente" "pendiente" "normal"    "patona"    "patona"
52
53 survey$yeti = ifelse(survey$yeti != "pendiente", survey$yeti, ifelse(survey$shoeSize >= 11,
54   "patona", ifelse(((survey$shoeSize < 11) & (survey$shoeSize >= 8)), "normal", "patita" )))
55
56 # resetear la variable poner todos pendientes
57 # survey$yeti = "pendiente"
58
59 survey[2,]
60
61 # Sigue estos ejemplos para ver si la variable es consistente
62
63 MujPatonas = survey[which(survey$yeti == "patona" & survey$gender == "female", arr.ind = TRUE),]
64 dim(MujPatonas)
65 HomPatonas = survey[which(survey$yeti == "patona" & survey$gender == "male", arr.ind = TRUE),]
66 dim(HomPatonas)
67 HomPatonas
68
69 ## Haz para patitas
70
```

Si es hombre queda "pendiente"

No es "pendiente" es TRUE

Queda igual

Paso 5 Graficar



```
74 library(ggplot2)
75
76 ggplot(data = survey, aes(x = yeti, fill = gender)) +
77   geom_bar() +
78   ylab("Personas") +
79   xlab("Tamaño del pie")+
80   geom_text(stat='count', aes(label=..count..), position = position_stack(vjust = 0.5))+
81   scale_fill_manual(values = c("pink", "#5bc4b9"), name = "Género", labels = c("Femenino", "Masculino"))
82
```