PEC2 Visualizacion 3

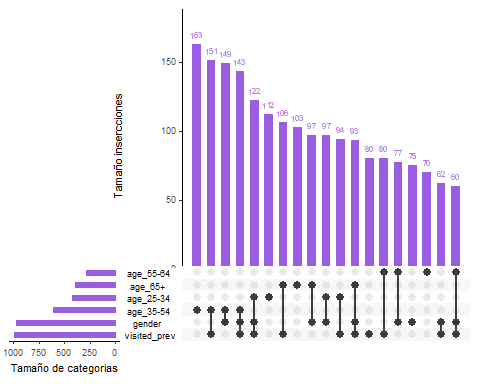
sonia rodriguez del cerro garcia

noviembre de 2021

# Cargamos los paquetes R que vamos a usar  
library(UpSetR)  
library(mltools)  
library(data.table)  
library(tidyverse)  
# Cargamos los datos  
set.seed(69)  
resp\_id <- 1:2000  
gender <- round(runif(2000,min=0,max=1)) #0 is male, 1 is female  
age <- sample(c("18-24","25-34","35-54","55-64","65+"),2000,replace=TRUE,prob=c(0.15,0.20,0.3,0.15,0.2))  
visited\_prev <- round(runif(2000,min=0,max=1)) #Whether respondent has visited before  
csat <- round(runif(2000,min=0,max=10)) #Customer Satisfaction Score  
csat[which(visited\_prev==0)] <- NA #NA if respondent had not visited before  
df <- data.frame(resp\_id,gender,age,visited\_prev,csat) #assign to a data frame

newdata <- one\_hot (as.data.table (df))

upset(newdata, nsets = 6,order.by = 'freq', main.bar.color = '#995ee1', sets.bar.color = '#995ee1', group.by = 'degree',mainbar.y.label = "Tamaño insercciones", sets.x.label = "Tamaño de categorias",number.angles = 5, point.size = 2.5, line.size = 1)



\*\*\*\*\*\*  
# Interpretación  
\*\*\*\*\*\*  
## Base de datos  
Se han generado datos ficticios de una encuesta con 2000 muestras, estaremos simulando nuestros propios datos de encuestas en lugar de importar un conjunto de datos previamente preparado.   
## Barras laterales  
Representan el tamaño total de cada categoría.   
En el caso de Gender hay como 1000 de los casos donde el genero el masculino.  
## Barras verticales  
Representa el tamaño de cada categoría o conjunto que se cruza. Se se completa con un punto negro, significa que esa categoría se incluye en el conjunto.  
## Ejemplo  
El primer conjunto nos indica que 151 personas de la encuesta habían realizado una visita anteriormente (visited\_prev) y su edad está entre 35 y 54 años.  
## Conclusión   
Si observamos que el segundo y cuarto grupo podemos ver se trantan de personas que habian realizado una visita anteriormente y su edad está entre 35 y 54 años. En el caso de el segundo grupo son hombres (gender = 0) y en el caso del 4 grupo son mujeres (gender = 1). Si sumamos ambos grupos 141 + 143 = 284 personas, lo que supone un 14% de los visitantes que cumplen esas caracteristicas.