

Tarea 1 - EYP2706

Natalie Julian

Septiembre

Actividad 1

La base de datos *hotel* contiene información sobre reservaciones de distintos hoteles. Interesa estudiar la variable target *Cancel* la que indica si la reservación fue o no cancelada.

- **Cancel:** Indica si la reserva fue cancelada o no (1: Cancelada, 0: No cancelada)
- **Mes:** Mes de la reserva
- **Weekend:** Cantidad de días de fin de semana (Sábados o Domingos) reservados
- **Weekday:** Cantidad de días de la semana (Lunes, Martes, Miércoles, Jueves, Viernes) reservados

```
library(readxl)
hotel <- read_excel("hotel.xlsx") #Carga la base de datos

#install.packages("tidyverse")
library(tidyverse)

glimpse(hotel) #Tipo de variables en la base de datos hotel
```

```
# Rows: 118,987
# Columns: 4
# $ Cancel <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0,...
# $ Mes <chr> "July", "July", "July", "July", "July", "July", "July", "July", "Ju...
# $ Weekend <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
# $ Weekday <dbl> 0, 0, 1, 1, 2, 2, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 1, 1, 4,...
```

```
summary(hotel) #Análisis descriptivo de las variables cuantitativas
```

#	Cancel	Mes	Weekend	Weekday
# Min.	:0.0000	Length:118987	Min. : 0.000	Min. : 0.000
# 1st Qu.:	:0.0000	Class :character	1st Qu.: 0.000	1st Qu.: 1.000
# Median	:0.0000	Mode :character	Median : 1.000	Median : 2.000
# Mean	:0.3708		Mean : 0.927	Mean : 2.499
# 3rd Qu.:	:1.0000		3rd Qu.: 2.000	3rd Qu.: 3.000
# Max.	:1.0000		Max. :19.000	Max. :50.000

Realizamos un gráfico:

