

EDAD EN BANCA

Erick Marroquín

2026-01-26

Análisis de promo “Adulto Mayor con Pisto”

Exploracion

```
library(nortest)

churn <- read.csv("Churn_Modelling.csv")

str(churn)

## 'data.frame': 10000 obs. of 14 variables:
## $ RowNumber : int 1 2 3 4 5 6 7 8 9 10 ...
## $ CustomerId : int 15634602 15647311 15619304 15701354 15737888 15574012 15592531 15656148 157...
## $ Surname   : chr "Hargrave" "Hill" "Onio" "Boni" ...
## $ CreditScore: int 619 608 502 699 850 645 822 376 501 684 ...
## $ Geography  : chr "France" "Spain" "France" "France" ...
## $ Gender    : chr "Female" "Female" "Female" "Female" ...
## $ Age       : int 42 41 42 39 43 44 50 29 44 27 ...
## $ Tenure    : int 2 1 8 1 2 8 7 4 4 2 ...
## $ Balance   : num 0 83808 159661 0 125511 ...
## $ NumOfProducts: int 1 1 3 2 1 2 2 4 2 1 ...
## $ HasCrCard : int 1 0 1 0 1 1 1 1 0 1 ...
## $ IsActiveMember: int 1 1 0 0 1 0 1 0 1 1 ...
## $ EstimatedSalary: num 101349 112543 113932 93827 79084 ...
## $ Exited    : int 1 0 1 0 0 1 0 1 0 0 ...

summary(churn)

##      RowNumber      CustomerId      Surname      CreditScore
## Min.   : 1   Min.   :15565701   Length:10000   Min.   :350.0
## 1st Qu.: 2501  1st Qu.:15628528   Class  :character  1st Qu.:584.0
## Median : 5000  Median :15690738   Mode   :character  Median :652.0
## Mean   : 5000  Mean   :15690941                    Mean   :650.5
## 3rd Qu.: 7500  3rd Qu.:15753234                    3rd Qu.:718.0
## Max.   :10000  Max.   :15815690                    Max.   :850.0
## 
##      Geography      Gender      Age      Tenure
## Length:10000  Length:10000  Min.   :18.00  Min.   : 0.000
## Class  :character  Class  :character  1st Qu.:32.00  1st Qu.: 3.000
## Mode   :character  Mode   :character  Median :37.00  Median : 5.000
## 
## 
## 
##      NumOfProducts  HasCrCard  IsActiveMember
## 
##      Balance
```

```

##  Min.   :    0   Min.   :1.00   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:    0   1st Qu.:1.00   1st Qu.:0.0000   1st Qu.:0.0000
##  Median : 97199   Median :1.00   Median :1.0000   Median :1.0000
##  Mean   : 76486   Mean   :1.53   Mean   :0.7055   Mean   :0.5151
##  3rd Qu.:127644   3rd Qu.:2.00   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :250898   Max.   :4.00   Max.   :1.0000   Max.   :1.0000
##  EstimatedSalary      Exited
##  Min.   : 11.58   Min.   :0.0000
##  1st Qu.: 51002.11  1st Qu.:0.0000
##  Median :100193.91  Median :0.0000
##  Mean   :100090.24  Mean   :0.2037
##  3rd Qu.:149388.25  3rd Qu.:0.0000
##  Max.   :199992.48  Max.   :1.0000

nrow(churn)

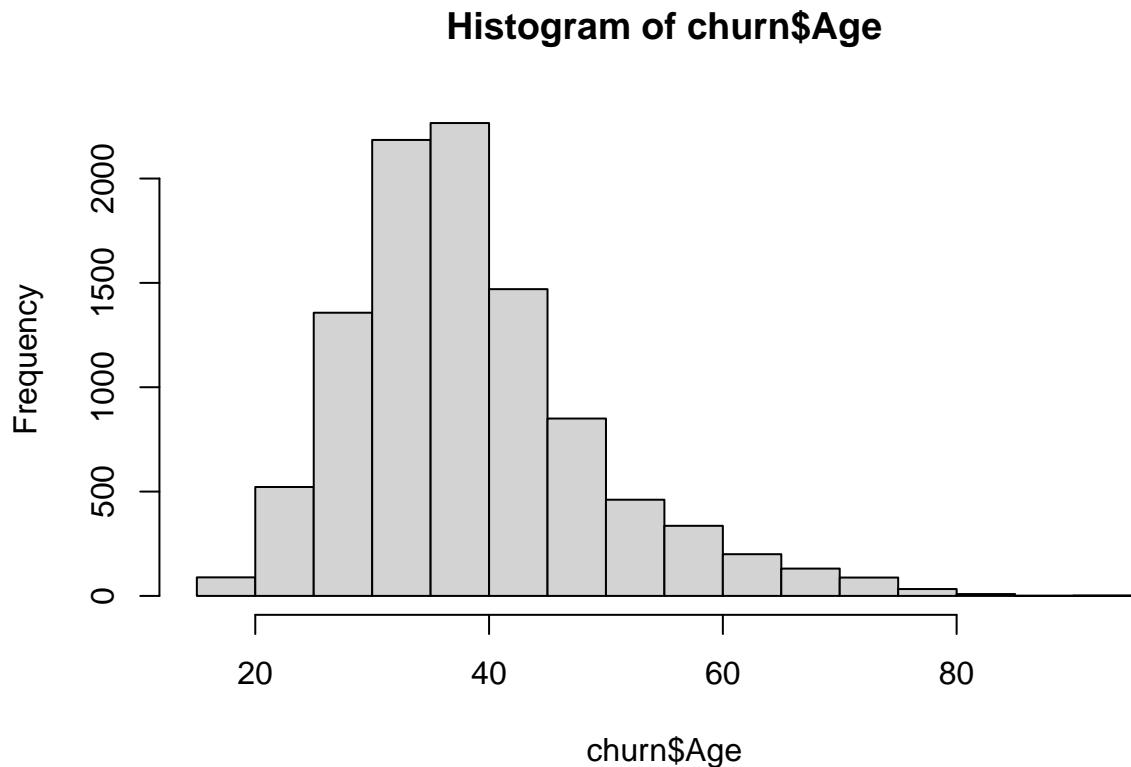
```

```
## [1] 10000
```

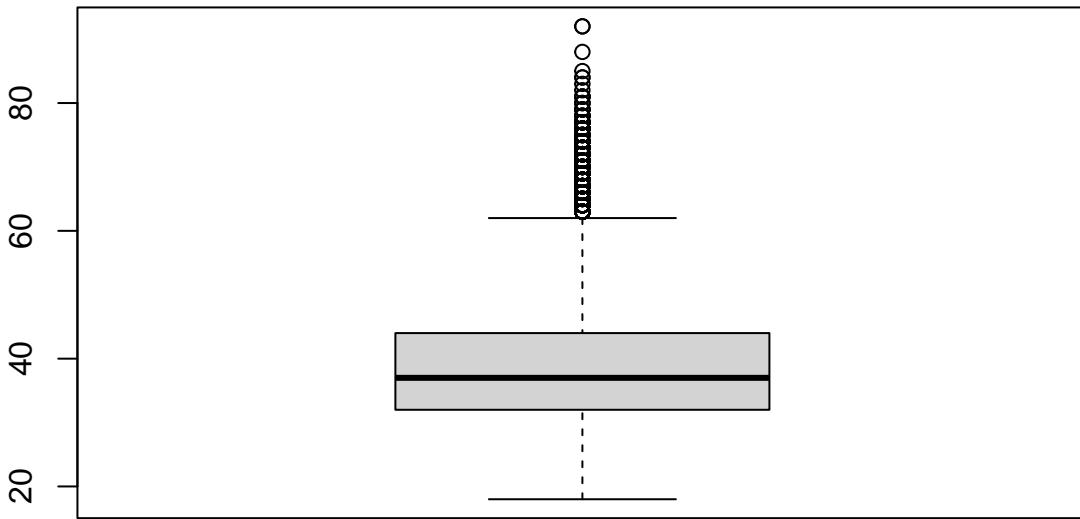
Como podemos observar la edad (en años) está en el tipo de datos correcto)

Graficos

```
hist(churn$Age)
```



```
boxplot(churn$Age)
```



Es evidente que la edad NO tiene una distribución normal. Gráficamente lo hemos confirmado pero es necesario una prueba de lilliforce para determinar esto con exactitud.

Prueba Lilliforce

```

cuant_discretas <- c(
  "CreditScore",
  "Tenure",
  "NumOfProducts"
)

cuant_continuas <- c(
  "Age",
  "Balance",
  "EstimatedSalary"
)

cuantitativas <- c(cuant_discretas, cuant_continuas)

datos_cuant <- churn[, cuantitativas]

lillie_results <- lapply(datos_cuant, lillie.test)

lillie_results$Age

##

```

```
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data: X[[i]]
## D = 0.10096, p-value < 2.2e-16
```

Como podemos ver el valor p es mucho menor al criterio de 0.05 por lo que aceptamos la H1 y concuimos que la data no tienen una distribución normal.