Campos NA - Missing Values

Jose R. Guignan

2023-08-11

```
#La biblioteca MICE es especializada para NA

nba <- read.csv("nba.csv")
head(nba)
```

```
##            Name           Team Number Position Age   Height Weight
## 1 Avery Bradley Boston Celtics      0       PG  25 170.9761    180
## 2   Jae Crowder Boston Celtics     99       SF  25 202.1928    235
## 3  John Holland Boston Celtics     30       SG  27 201.3758    205
## 4   R.J. Hunter Boston Celtics     28       SG  22 181.7122    185
## 5 Jonas Jerebko Boston Celtics      8       PF  29 190.5045    231
## 6  Amir Johnson Boston Celtics     90       PF  29 215.2288    240
##             College   Salary
## 1             Texas  7730337
## 2         Marquette  6796117
## 3 Boston University       NA
## 4     Georgia State  1148640
## 5                    5000000
## 6                   12000000
```

```
#Contar NA
vector = c(2,NA,5,8,10,NA,4,NA)
is.na(vector)
```

```
## [1] FALSE  TRUE FALSE FALSE FALSE  TRUE FALSE  TRUE
```

```
#Numeros de Na - TRUE vale 1 y FALSE 0, en numerico
sum(is.na(vector))
```

```
## [1] 3
```

```
#metodo 1
sum(is.na(nba$Salary))
```

```
## [1] 11
```

```
#metodo 2
nrow(nba[is.na(nba$Salary),])
```

```
## [1] 11
```

```r
#metodo 3
apply(nba, MARGIN = 2, function(x) sum(is.na(x)))
```

```
##     Name     Team   Number Position      Age   Height   Weight  College
##        0        0        0        0        0        0        0        0
##   Salary
##       11
```

```r
#Coloca el valor de la mediana, donde estan los NA
#na.rm=TRUE se debe colocar para que calcule la mediana sin tener en cuenta los NA
nba[is.na(nba$Salary),"Salary"] <- median(nba$Salary, na.rm = TRUE)

sum(is.na(nba$Salary))
```

```
## [1] 0
```