

Recodificación de variables

Irais Espejo

2022-07-04

RECODIFICACIÓN DE VARIABLES

Se trabaja con la matriz “Escuela”

1. Importación de la matriz

```
library(readxl)
Escuela <- read_excel("Escuela.xlsx")
```

Exploración de la matriz

1. Número de variables

```
dim(Escuela)
```

```
## [1] 59 7
```

2. Organización basada en factores

```
str(Escuela)
```

```
## tibble [59 x 7] (S3: tbl_df/tbl/data.frame)
## $ ID      : chr [1:59] "i1" "i2" "i3" "i4" ...
## $ grupo   : chr [1:59] "A" "A" "A" "A" ...
## $ genero   : chr [1:59] "Femenino" "Femenino" "Femenino" "Femenino" ...
## $ Matematicas: num [1:59] 7 8 9 9 8 9 10 9 8 10 ...
## $ Español  : num [1:59] 8 9 9 9 9 9 9 8 8 9 ...
## $ Ciencias : num [1:59] 8 10 8 9 9 9 9 9 9 9 ...
## $ Geografia : num [1:59] 8 10 10 9 9 9 9 9 9 9 ...
```

3. Presentación de datos: calificación mínima, máxima. Mediana. primer y tercer cuartil

```
summary(Escuela)
```

```
##      ID              grupo          genero      Matematicas
## Length:59      Length:59      Length:59      Min.    : 7.000
## Class :character Class :character Class :character 1st Qu.: 8.000
```

```
## Mode :character Mode :character Mode :character Median : 9.000
## Mean : 8.576
## 3rd Qu.: 9.000
## Max. :10.000
## Español Ciencias Geografia
## Min. : 8.000 Min. : 7.000 Min. : 8.000
## 1st Qu.: 9.000 1st Qu.: 8.000 1st Qu.: 9.000
## Median : 9.000 Median : 9.000 Median : 9.000
## Mean : 8.932 Mean : 8.492 Mean : 9.119
## 3rd Qu.: 9.000 3rd Qu.: 9.000 3rd Qu.: 9.500
## Max. :10.000 Max. :10.000 Max. :10.000
```

Configuración de variables

```
Escuela$genero<-factor(Escuela$genero,
                        levels=c("Femenino", "Masculino"))
```

```
Escuela$Matematicas<-factor(Escuela$Matematicas,
                             levels=c("5","6","7","8","9","10"))
```

```
Escuela$Español<-factor(Escuela$Español,
                         levels=c("5","6","7","8","9","10"))
```

```
Escuela$Ciencias<-factor(Escuela$Ciencias,
                          levels=c("5","6","7","8","9","10"))
```

```
Escuela$Geografia<-factor(Escuela$Geografia,
                           levels=c("5","6","7","8","9","10"))
```

```
dim(Escuela)
```

```
## [1] 59 7
```

```
colnames(Escuela)
```

```
## [1] "ID" "grupo" "genero" "Matematicas" "Español"
## [6] "Ciencias" "Geografia"
```

```
str(Escuela)
```

```
## tibble [59 x 7] (S3: tbl_df/tbl/data.frame)
## $ ID : chr [1:59] "i1" "i2" "i3" "i4" ...
## $ grupo : chr [1:59] "A" "A" "A" "A" ...
## $ genero : Factor w/ 2 levels "Femenino","Masculino": 1 1 1 1 1 2 2 2 2 2 ...
## $ Matematicas: Factor w/ 6 levels "5","6","7","8",...: 3 4 5 5 4 5 6 5 4 6 ...
## $ Español : Factor w/ 6 levels "5","6","7","8",...: 4 5 5 5 5 5 5 4 4 5 ...
## $ Ciencias : Factor w/ 6 levels "5","6","7","8",...: 4 6 4 5 5 5 5 5 5 5 ...
## $ Geografia : Factor w/ 6 levels "5","6","7","8",...: 4 6 6 5 5 5 5 5 5 5 ...
```

```
anyNA(Escuela)
```

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

```
summary(Escuela)
```

```
##      ID              grupo      genero  Matematicas Español
## Length:59      Length:59      Femenino :20    5 : 0      5 : 0
## Class :character Class :character Masculino:39    6 : 0      6 : 0
## Mode  :character Mode  :character      7 : 7      7 : 0
##                                     8 :19      8 :12
##                                     9 :25      9 :39
##                                     10: 8      10: 8
## Ciencias Geografia
## 5 : 0      5 : 0
## 6 : 0      6 : 0
## 7 : 9      7 : 0
## 8 :13      8 : 8
## 9 :36      9 :36
## 10: 1      10:15
```

Gráficos

Abrir librería ggplot2

```
library(ggplot2)
```

1. Creación de un vector de color

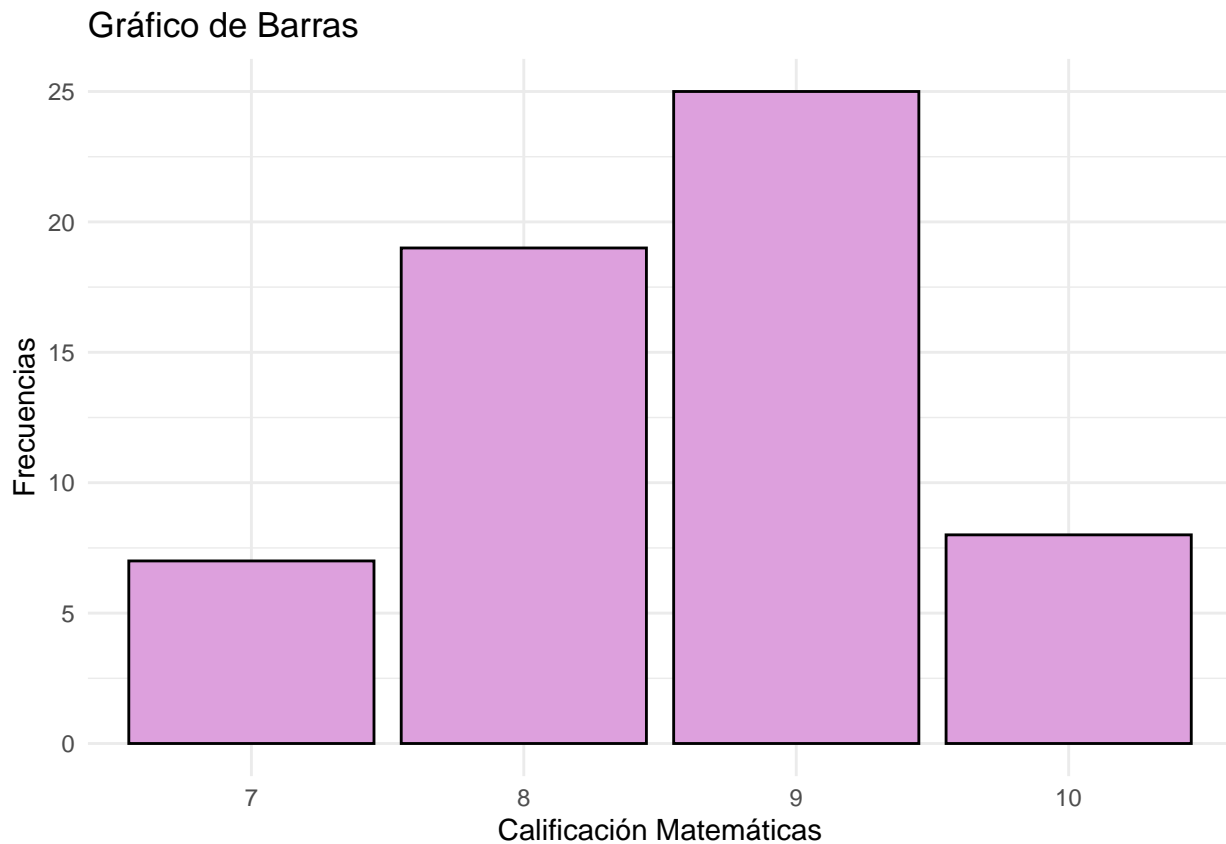
```
color=c("plum")
```

2. Creación del gráfico

```
GB1<-ggplot(Escuela, aes(x=Matematicas))+
  geom_bar(colour= "black", fill=color)+
  ggtitle("Gráfico de Barras")+
  xlab("Calificación Matemáticas")+
  ylab("Frecuencias")+
  theme_minimal()
```

3. Visualización del gráfico

GB1



1. Creación de un vector de color

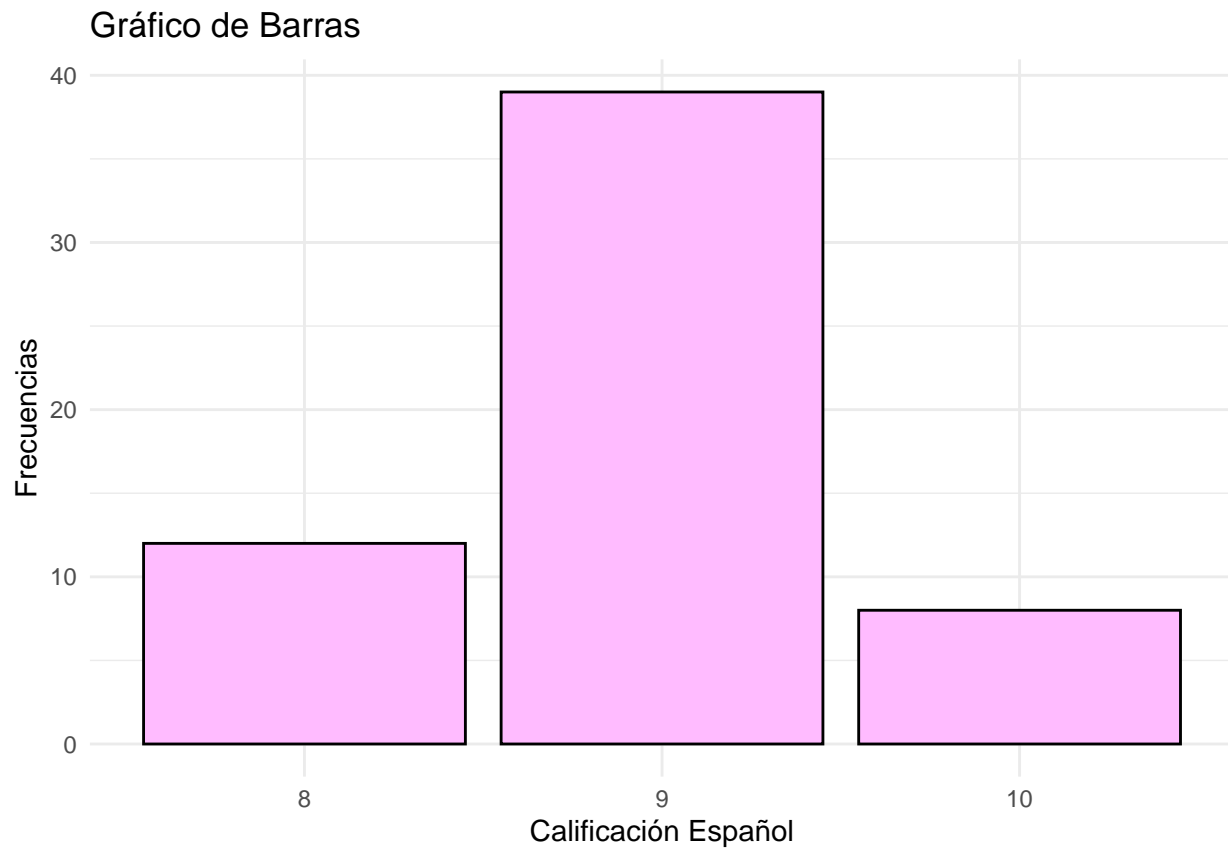
```
color=c("plum1")
```

2. Creación del gráfico

```
GB2<-ggplot(Escuela, aes(x=Español))+  
  geom_bar(colour= "black", fill=color)+  
  ggtitle("Gráfico de Barras")+  
  xlab("Calificación Español")+  
  ylab("Frecuencias")+  
  theme_minimal()
```

3. Visualización del gráfico

GB2



1. Creación de un vector de color

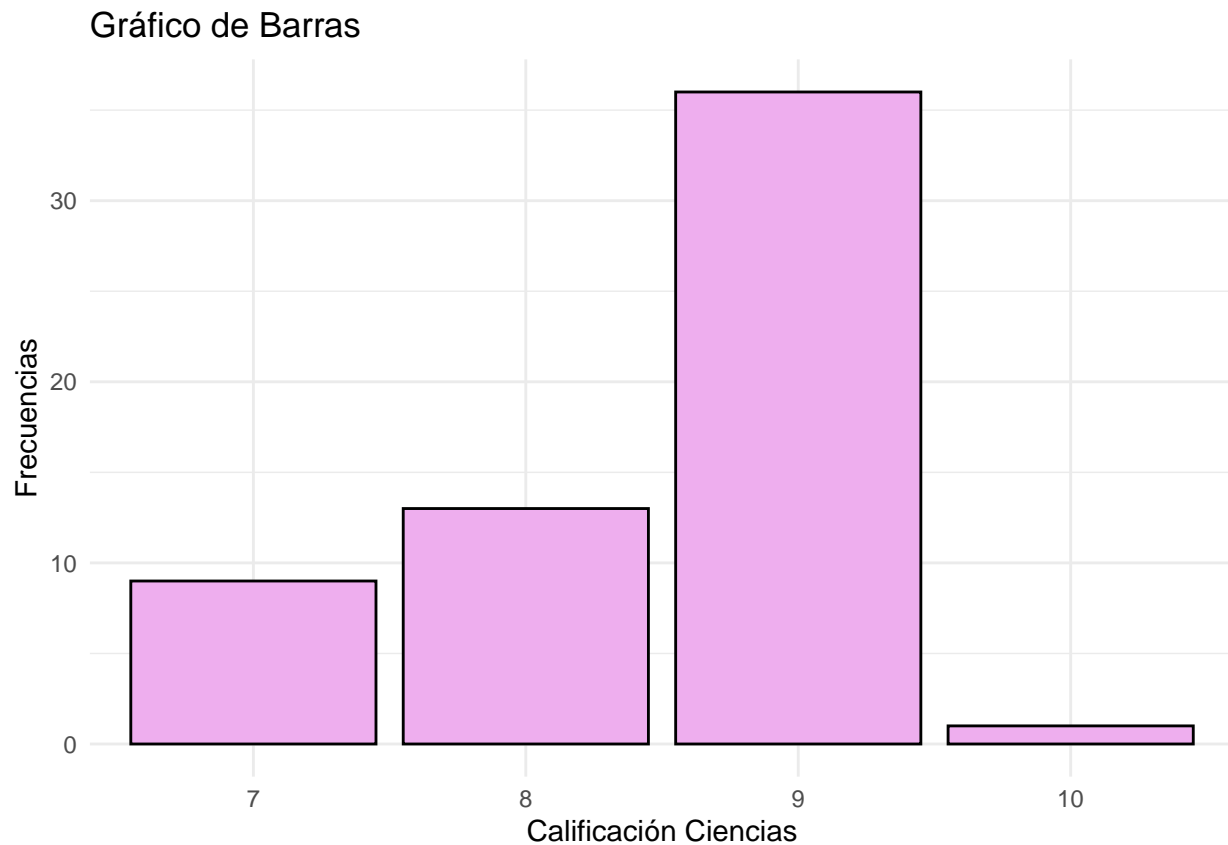
```
color=c("plum2")
```

2. Creación del gráfico

```
GB3<-ggplot(Escuela, aes(x=Ciencias))+  
  geom_bar(colour= "black", fill=color)+  
  ggtitle("Gráfico de Barras")+  
  xlab("Calificación Ciencias")+  
  ylab("Frecuencias")+  
  theme_minimal()
```

3. Visualización del gráfico

```
GB3
```



1. Creación de un vector de color

```
color=c("plum3")
```

2. Creación del gráfico

```
GB4<-ggplot(Escuela, aes(x=Geografia))+  
  geom_bar(colour= "black", fill=color)+  
  ggtitle("Gráfico de Barras")+  
  xlab("Calificación Geografía")+  
  ylab("Frecuencias")+  
  theme_minimal()
```

3. Visualización del gráfico

```
GB4
```

