

# Fatores e fatores ordenados

Ivanildo Batista

15 de março de 2021

## FATORES E FATORES ORDENADOS

Serve para armazenar variáveis categóricas (nominais e ordinais) e usa-se a função `factor()`

Criando vetor

```
vec1 = c('Macho', 'Femea', 'Femea', 'Macho', 'Macho')
```

```
vec1
```

```
## [1] "Macho" "Femea" "Femea" "Macho" "Macho"
```

```
fac_vec1 = factor(vec1) #criando um fator com o vetor criado  
fac_vec1
```

```
## [1] Macho Femea Femea Macho Macho  
## Levels: Femea Macho
```

```
class(vec1) #verificando classe do vetor
```

```
## [1] "character"
```

```
class(fac_vec1) #verificando classe do fator
```

```
## [1] "factor"
```

Variáveis categóricas nominais (não existe uma ordem explícita)

```
animais = c('Zebra', 'Pantera', 'Rinoceronte', 'Macaco', 'Tigre')  
animais
```

```
## [1] "Zebra"      "Pantera"    "Rinoceronte" "Macaco"     "Tigre"
```

```
class(animais) #verificando a classe
```

```
## [1] "character"
```

```
fac_animais = factor(animais)#transformando em um fator
fac_animais
```

```
## [1] Zebra      Pantera      Rinoceronte Macaco      Tigre
## Levels: Macaco Pantera Rinoceronte Tigre Zebra
```

```
class(fac_animais) #verificando a classe
```

```
## [1] "factor"
```

```
levels(fac_animais) #verificando os níveis do fator
```

```
## [1] "Macaco"      "Pantera"      "Rinoceronte" "Tigre"        "Zebra"
```

### Variáveis categóricas ordinais

```
grad = c('Mestrado','Doutorado','Bacharelado','Mestrado','Mestrado')
grad
```

```
## [1] "Mestrado"      "Doutorado"      "Bacharelado" "Mestrado"      "Mestrado"
```

```
fac_grad = factor(grad, order=TRUE,
                  levels = c('Doutorado','Mestrado','Bacharelado'))

summary(fac_grad) #sumarizando os dados do fator ordenado
```

```
##      Doutorado      Mestrado Bacharelado
##           1           3           1
```

```
summary(grad) #sumarizando os dados do vetor
```

```
##      Length      Class      Mode
##           5 character character
```

### Outro exemplo

```
vec2 = c('M','F','F','M','M','M','F','F','M','M','F')
vec2
```

```
## [1] "M" "F" "F" "M" "M" "M" "F" "F" "M" "M" "F"
```

```
fac_vec2 = factor(vec2)
fac_vec2
```

```
## [1] M F F M M M F F M M F
## Levels: F M
```

```
levels(fac_vec2) = c('Femea','Macho') #atribuindo outros valores aos níveis do fator

fac_vec2 #verificando se os fatores foram alterados
```

```
## [1] Macho Femea Femea Macho Macho Macho Femea Femea Macho Macho Femea
## Levels: Femea Macho
```

## Mais exemplos

```
data = c(1,2,2,3,1,2,3,3,1,2,3,3,1) #criando um vetor
fdata = factor(data) #criando um fator com o vetor
fdata
```

```
## [1] 1 2 2 3 1 2 3 3 1 2 3 3 1
## Levels: 1 2 3
```

```
rdata = factor(data, labels = c('I','II','III'))
rdata
```

```
## [1] I   II  II  III I   II  III III I   II  III III I
## Levels: I II III
```

## Fatores não ordenados

```
set1 = c('AA','B','BA','CC','CA','AA','BA','CC','CC','BA')
set1
```

```
## [1] "AA" "B"  "BA" "CC" "CA" "AA" "BA" "CC" "CC" "BA"
```

```
f.set1=factor(set1)
f.set1
```

```
## [1] AA B  BA CC CA AA BA CC CC BA
## Levels: AA B BA CA CC
```

```
class(f.set1)
```

```
## [1] "factor"
```

```
is.ordered(f.set1) #verificando se o fator está ordenado
```

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

## Fatores ordenados

```
#ordenando o fator
o.set1 = factor(set1, levels =c('CA','BA','AA','CC','B'),ordered = TRUE)
o.set1
```

```
## [1] AA B BA CC CA AA BA CC CC BA
## Levels: CA < BA < AA < CC < B
```

```
is.ordered(o.set1)
```

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

```
as.numeric(o.set1)# convertendo para valor numérico
```

```
## [1] 3 5 2 4 1 3 2 4 4 2
```

```
table(o.set1)
```

```
## o.set1
## CA BA AA CC B
## 1 3 2 3 1
```

## Fatores e Dataframes

```
df = read.csv2('C:/Users/junio/etnias.csv',sep=',')
df
```

```
##      Codigo Sexo Idade Peso Altura Estado_Civil  Etnia
## 1    1000    M    35   89   175              0 Branco
## 2    1001    F    23   70   160              1 Pardo
## 3    1002    F    34   64   161              1 Pardo
## 4    1003    F    76   58   167              2 Negro
## 5    1004    M    58   56   159              0 Branco
## 6    1005    M    43   63   167              0 Branco
## 7    1006    F    42   62   172              2 Negro
## 8    1007    M    29  120   190              1 Pardo
## 9    1008    F    23   88   189              1 Pardo
## 10   1009    M    67   81   186              1 Branco
## 11   1010    M    58   64   175              0 Branco
```

```
str(df) #informações sobre o dataframe
```

```
## 'data.frame':   11 obs. of  7 variables:
## $ Codigo      : int  1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 ...
## $ Sexo       : chr   "M" "F" "F" "F" ...
## $ Idade      : int   35 23 34 76 58 43 42 29 23 67 ...
## $ Peso       : int   89 70 64 58 56 63 62 120 88 81 ...
## $ Altura     : int  175 160 161 167 159 167 172 190 189 186 ...
## $ Estado_Civil: int   0 1 1 2 0 0 2 1 1 1 ...
## $ Etnia      : chr   "Branco" "Pardo" "Pardo" "Negro" ...
```

```
levels(df$Etnia)#níveis da coluna Etnia
```

```
## NULL
```

```
df$Etnia =factor(df$Etnia)
```

```
str(df) #novamente informações sobre o dataframe
```

```
## 'data.frame':   11 obs. of  7 variables:
## $ Codigo      : int  1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 ...
## $ Sexo       : chr   "M" "F" "F" "F" ...
## $ Idade      : int   35 23 34 76 58 43 42 29 23 67 ...
## $ Peso       : int   89 70 64 58 56 63 62 120 88 81 ...
## $ Altura     : int  175 160 161 167 159 167 172 190 189 186 ...
## $ Estado_Civil: int   0 1 1 2 0 0 2 1 1 1 ...
## $ Etnia      : Factor w/ 3 levels "Branco","Negro",...: 1 3 3 2 1 1 2 3 3 1 ...
```

```
levels(df$Etnia)#níveis da coluna Etnia
```

```
## [1] "Branco" "Negro" "Pardo"
```

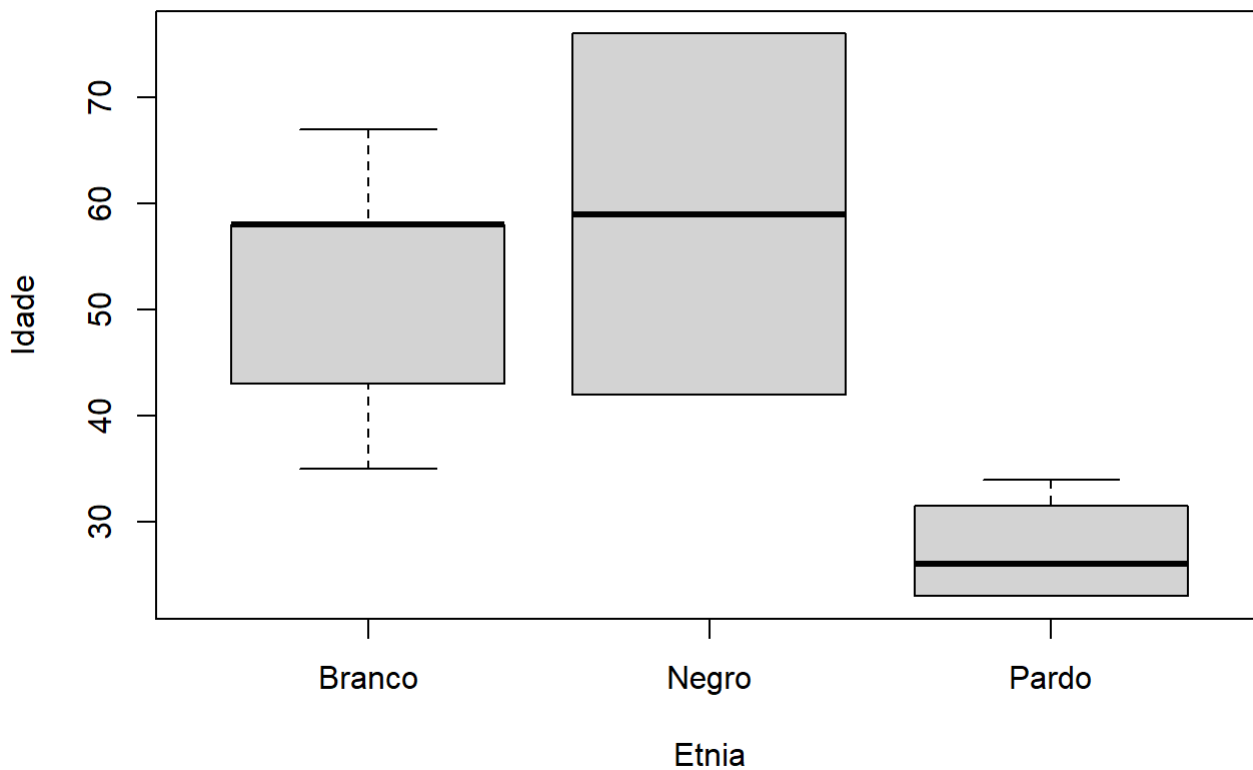
```
summary(df$Etnia)
```

```
## Branco  Negro  Pardo
##       5      2      4
```

Plotando gráfico com colunas do dataframe

```
plot(df$Idade~df$Etnia, xlab = 'Etnia',ylab = 'Idade',
     main = 'Idade por Etnia')
```

## Idade por Etnia



Gerando uma regressão com as colunas do dataframe

```
summary(lm(Idade~Etnia, data = df))
```

```
##
## Call:
## lm(formula = Idade ~ Etnia, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.200  -6.725   1.750   6.275  17.000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   52.200     5.765   9.054 1.77e-05 ***
## EtniaNegro     6.800    10.786   0.630  0.5460
## EtniaPardo   -24.950     8.648  -2.885  0.0204 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.89 on 8 degrees of freedom
## Multiple R-squared:  0.5892, Adjusted R-squared:  0.4865
## F-statistic: 5.737 on 2 and 8 DF,  p-value: 0.02848
```

Convertendo uma coluna para fator

```
df$Estado_Civil.cat = factor(df$Estado_Civil,
                             labels = c('Solteiro','Casado','Divorciado'))
```

```
df
```

```
##      Codigo Sexo Idade Peso Altura Estado_Civil  Etnia Estado_Civil.cat
## 1    1000     M   35   89   175           0 Branco      Solteiro
## 2    1001     F   23   70   160           1 Pardo       Casado
## 3    1002     F   34   64   161           1 Pardo       Casado
## 4    1003     F   76   58   167           2 Negro       Divorciado
## 5    1004     M   58   56   159           0 Branco      Solteiro
## 6    1005     M   43   63   167           0 Branco      Solteiro
## 7    1006     F   42   62   172           2 Negro       Divorciado
## 8    1007     M   29  120   190           1 Pardo       Casado
## 9    1008     F   23   88   189           1 Pardo       Casado
## 10   1009     M   67   81   186           1 Branco      Casado
## 11   1010     M   58   64   175           0 Branco      Solteiro
```

```
str(df)
```

```
## 'data.frame':    11 obs. of  8 variables:
## $ Codigo      : int  1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 ...
## $ Sexo        : chr  "M" "F" "F" "F" ...
## $ Idade        : int  35 23 34 76 58 43 42 29 23 67 ...
## $ Peso         : int  89 70 64 58 56 63 62 120 88 81 ...
## $ Altura       : int  175 160 161 167 159 167 172 190 189 186 ...
## $ Estado_Civil : int   0 1 1 2 0 0 2 1 1 1 ...
## $ Etnia        : Factor w/ 3 levels "Branco","Negro",...: 1 3 3 2 1 1 2 3 3 1 ...
## $ Estado_Civil.cat: Factor w/ 3 levels "Solteiro","Casado",...: 1 2 2 3 1 1 3 2 2 2 ...
```