

Curso Livre I

Aluno: Manuel Ferreira Junior

Matricula: 20180008601

Modulo I

Tarefa 02

A) Crie um vetor para cada variável do banco de dados

```
salario <-  
c(249420000.00, 230555000.00, 139960000.00, 135530000.00, 122670000.00,  
   80730000.00, 75330000.00, 71840000.00, 69660000.00, 68955000.00,  
   62990000.00, 56470000.00, 55990000.00, 47830000.00, 46190000.00,  
   41310000.00, 39635000.00, 39315000.00, 37480000.00, 37420000.00,  
   36540000.00, 36380000.00, 35410000.00, 34490000.00, 32935000.00,  
   32730000.00, 32190000.00, 31540000.00, 31436000.00, 31435000.00,  
   31340000.00, 31230000.00, 30866000.00, 29360000.00, 28820000.00,  
   28400000.00, 27970000.00, 27885000.00, 26900000.00, 25180000.00)  
  
idade <- c(55, 63, 65, 60, 60, 71, 61, 57, 57, 67, 56, 58, 60, 48, 62, 65,  
           63, 53, 55, 53, 59, 60, 59, 58, 43, 64, 57, 63, 64, 63, 59, 61,  
           57, 53, 54, 72, 54, 50, 60, 62)  
  
formacao <- c(2, 2, 3, 3, 2, 5, 0, 4, 1, 1, 2, 1, 2, 2, 4, 3, 1, 2, 1, 1, 2, 2, 2,  
              4, 2, 2, 2, 2, 5, 2, 2, 1, 4, 2, 2, 0, 1, 2, 3, 2)
```

B) Divida a variável salário por um milhão

```
salario <- salario/1e6 # Armazenando dentro de salario
```

C) Transforma a variável formação em um fator, em que 0 – nenhuma, 1 – bacharelado, 2 – MBA, 3 – LLM, 4 – Mestrado, 5 – PhD

```
factors <- c('nenhum', 'Bacharelado', 'MBA', 'LLM', 'Mestrado', 'PhD') #  
Criando um vetor indicando os fatores  
  
formacao <- factor(formacao,  
                   levels=min(formacao):max(formacao),  
                   labels=factors) # Criando um factor
```

D) Crie um data frame com a variável idade e com as novas variáveis dos itens B) e C).

```
df <- data.frame('salario' = salario, 'formacao'=formacao)
# Criando um dataframe com salario e formacao, itens usados em B e C
```

E) Calcule o valor máximo e o mínimo da variável salário, e o tamanho da amostra.

```
> c('minimo (salario)'=min(df$salario), 'maximo (salario)'=max(df$salario))
minimo (salario) maximo (salario)
      25.18      249.42
> # Vetor mostrando o minimo e o maximo salario
> #n <- nrow(df) ou
> n <- length(df$salario)
> n # Tamanho da amostra
[1] 40
```

F) Faça o mesmo para a variavel idade

```
> c('minimo (idade)'=min(idade), 'maximo (idade)'=max(idade))
minimo (idade) maximo (idade)
      43      72
> # Vetor mostrando a menor, e a maior idade
```

G) Classifique a variável obtida na letra b) em dois fatores, sendo Fator 1 os salários menores ou iguais a 369,8 milhões de dólares e o Fator 2 os salários maiores que 36,98

```
fact.sal <- ifelse(salario <= 36.98, 1, 2)
# Fiz usando 36.98, pois achei que havia algum erro no enunciado da
pergunta, uma ideia ambigua.
# criando um vetor de "1" e "2", caso o salario seja inferior ou igual a
36.98 dolares, retorne 1, caso superior retorne 2
```

H) Obtenha o tamanho de cada fator da letra G)

```
> c('salario <= 36.98'=sum(fact.sal == 1),
+   'salario > 36.98'=sum(fact.sal == 2))
salario <= 36.98  salario > 36.98
      20      20
> # Calculando quantos possuem salario inferior ou igual a 36.98, e quantos
possuem salario superior a 36.98
> # ou, simplesmente :
> # table(fact.sal)
```