

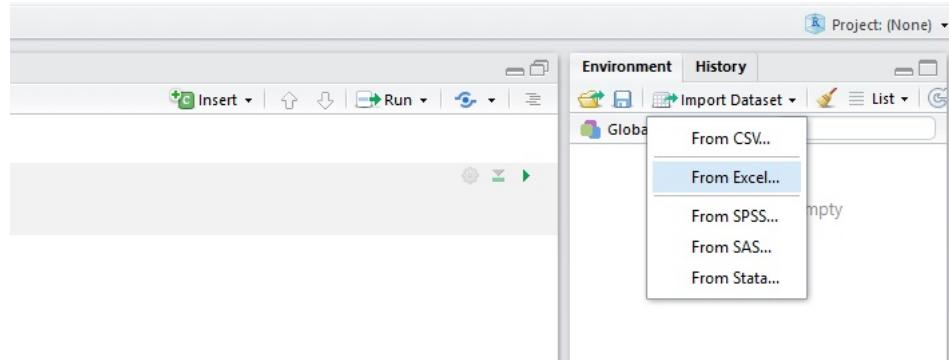
09

Faça o que eu fiz na aula

Primeiramente, você precisará importar a base “funcionários.xls” para o R. E também da `library(readxl)`. Caso não esteja instalada, instale utilizando o comando `install` e

```
library(readxl)
```

A importação da base no R pode ser feita pelo “Import Dataset” (no menu superior direito).



Selecione o arquivo na pasta de origem. Outra forma de importar o arquivo é diretamente pela linha de comando, utilizando:

```
funcionarios <- read_excel("caminho da base/funcionarios.xls")
View(funcionarios)
```

Em seguida, precisamos avaliar a estrutura da base de dados, utilizamos agora o comando:

```
str(funcionarios)
```

```
Classes 'tbl_df', 'tbl' and 'data.frame':      36 obs. of  8 variables:
 $ numero      : num  1 2 3 4 5 6 7 8 9 10 ...
 $ civil        : num  1 2 2 1 1 2 1 1 2 1 ...
 $ instrucao   : num  1 1 1 2 1 1 1 2 2 ...
 $ N de Filhos: chr  "NA" "2" "NA" ...
 $ salario      : chr  "4.00" "4.56" "5.25" "5.73" ...
 $ ano          : num  26 32 36 20 40 28 41 43 34 23 ...
 $ mes          : num  3 10 5 10 7 0 0 4 10 6 ...
 $ regiao       : num  3 3 3 3 3 3 3 3 3 3 ...
```

Observa-se, que todos os campos são do tipo `num` (numérico), porém, pode-se notar que alguns campos não são numéricos, como `civil` (estado civil), `instrucao` (grau de instrução) e `regiao`. Vamos transformar esses campos para o tipo `factor` (fator). Precisamos categorizar esses campos, utilizando o comando `mutate()` da seguinte forma:

```
func_t <- funcionarios %>%
  mutate(
    civil = factor(civil, label = c("solteiro", "casado"), levels = 1:2),
    instrucao = factor(instrucao, labels = c("1ºgrau", "2ºgrau", "superior"), levels = 1:3),
    regiao = factor(regiao, labels = c("capital", "interior", "outros"), levels = c(1:3)))
```

```
)
str(func_t)
```

```
Classes 'tbl_df', 'tbl' and 'data.frame':      36 obs. of  9 variables:
 $ numero    : num  1 2 3 4 5 6 7 8 9 10 ...
 $ civil      : Factor w/ 2 levels "solteiro","casado": 1 2 2 1 1 2 1 1 2 1 ...
 $ instrucao  : num  1 1 1 2 1 1 1 2 2 ...
 $ N de Filhos: chr  "NA" "1" "2" "NA" ...
 $ salario    : chr  "4.00" "4.56" "5.25" "5.73" ...
 $ ano        : num  26 32 36 20 40 28 41 43 34 23 ...
 $ mes        : num  3 10 5 10 7 0 0 4 10 6 ...
 $ regiao     : Factor w/ 3 levels "capital","interior",...: 3 3 3 3 3 3 3 3 3 3 ...
 $ instrucao  : Factor w/ 3 levels "1ºgrau","2ºgrau",...: 1 1 1 2 1 1 1 2 2 ...
```

Observe que, agora, os campos foram alterados para o tipo `Factor` (fator) e seus respectivos níveis (categorias) foram alterados.

```
head (func_t)
```

numero	civil	instrucao	N de Filhos	salario	ano	mes	regiao	instrucao
<dbl>	<fctr>	<dbl>	<chr>	<chr>	<dbl>	<dbl>	<fctr>	<fctr>
1	solteiro	1	NA	4.00	26	3	outros	1ºgrau
2	casado	1	1	4.56	32	10	outros	1ºgrau
3	casado	1	2	5.25	36	5	outros	1ºgrau
4	solteiro	2	NA	5.73	20	10	outros	2ºgrau
5	solteiro	1	NA	6.26	40	7	outros	1ºgrau
6	casado	1	0	6.66	28	0	outros	1ºgrau

6 rows