

O código está grifado de cinza. Tudo o que estiver em verde tem uma hashtag na frente, o que indica ao R que se trata de um comentário (e portanto não é rodado como código). Comentários são fundamentais para que o código seja compreensível no futuro.

1) Abrir o software RStudio

2) Clicar em File → New File → R Markdown

3) Insira um bloco de código apertando ctrl + alt + i, digitando dentro dele os seguintes comandos, que instalarão esses pacotes entre parênteses:

```
``#{r}
install.packages("openxlsx")
install.packages("ggplot2")
install.packages("lattice")
``
```

4) Insira um novo bloco de código e rode esses comandos para ativar os pacotes nessa sessão:

```
``#{r}
library(openxlsx)
library(ggplot2)
library(lattice)
``
```

5) Novamente insira um bloco de códigos, e rode o seguinte (lembrando que a hashtag diz ao R que o que for escrito dali em diante é comentário, e não será rodado como código):

```
``{r}
```

```
dados_basicos <-  
read.xlsx("https://github.com/bfsg839/CED3024/raw/main/dados_EDA.xlsx", sheet=1 )
```

```
as.data.frame(dados_basicos) # Usamos a função as.data.frame para forçar os dados em um  
formato melhor
```

```
# Dê uma primeira olhada nos dados, rodando a linha abaixo:
```

```
View(dados_basicos)
```

```
# Veja que há alguns NAs, isto é, valores que não estão disponíveis nos dados originais,
```

```
# provavelmente porque não foram coletados. Com a linha abaixo, vamos substituir os NAs  
por zero.
```

```
dados_basicos[is.na(dados_basicos)] <- 0
```

```
# Veja se deu certo, rodando novamente a função View.
```

```
``
```

6) Insira um novo bloco e rode o seguinte:

```
``{r}
```

```
# Para ver alguns parâmetros básicos, rode o comando summary:
```

```
summary(dados_basicos)
```

```
``
```

7) Novo bloco

```
``{r}
```

```
# Fazendo gráficos em barras para grau de instrução, para ver
```

```
# melhor o comportamento dessa variável
```

```
educ <- dados_basicos$"Grau.de.Instrução"
```

```
ggplot(dados_basicos, aes(x=educ)) + geom_bar()
```

```
# Faça uma busca para descobrir como alterar o nome dos eixos, cores do gráfico etc.
```

```
# Para isso, sugiro procurar por 'ggplot título gráfico', 'ggplot axis title', etc.
```

```
# Segue um exemplo de resultado que pode ajudar:
```

```
# http://www.sthda.com/english/wiki/ggplot2-title-main-axis-and-legend-titles
```

```
``
```

8) Novo bloco

```
``{r}
```

para criar um histograma com frequências relativas, primeiro vamos criar o objeto "numerosfilhos". Ao invés de criar o objeto "educ" contendo o grau de instrução, modifique a linha abaixo para criar o objeto "numerosfilhos" contendo a variável "N.de.Filhos". (Atenção: o R diferencia entre CAIXA ALTA e caixa alta)

```
educ <- dados_basicos$"Grau.de.Instrução"
```

Agora, para criar o histograma de frequência relativa do número de filhos, use a função histogram, como consta neste link:

<https://www.geeksforgeeks.org/how-to-create-a-relative-frequency-histogram-in-r/>

```
``
```

9) Novo bloco

Para facilitar a visualização e compreensão dos dados, faça uma busca e veja o que fazer para multiplicar a variável Salário (que está aparecendo como múltiplos do salário mínimo) pelo salário mínimo atual (R\$ 1320). Dica: faça a seguinte pergunta ao ChatGPT: "Estou programando em linguagem R. Preciso criar uma nova variável, multiplicando uma coluna do meu banco de dados por um único número (1320)."