

```
#Estatística Descritiva
```

```
if(!require(dplyr)) install.packages("dplyr")
```

```
if(!require(ggplot2)) install.packages("ggplot2")
```

```
library(dplyr)
```

```
library(ggplot2)
```

Comecei colocando essas linhas de código como se tivesse importando as bibliotecas que sabia que seriam usadas.

```
setwd("C:/Users/dpits/Desktop/eu/Repositorios/R/")
```

em seguida não lembrava desse comando, então pesquisei e vi que ele é usado para “setar” a pasta que vai ser trabalhada, isso facilita o acesso aos arquivos que serão usados, no caso aqui, somente o ds\_salaries.csv.

```
dataFrame <- read.csv("ds_salaries.csv", sep=",")
```

criei uma variável que vai representar o dataframe, o mesmo recebe as informações do CSV

```
View(dataFrame)
```

comando usado para ver o dataframe

```
excluir = c("company_size", "experience_level", 'salary',
```

```
'salary_currency', 'remote_ratio', "work_year",
```

```
"employment_type", 'employee_residence', 'company_location')
```

```
dataFrame = dataFrame[ , !(names(dataFrame) %in% excluir)]
```

```
View(dataFrame)
```

nesse outro momento é criada uma variável para auxiliar na exclusão das tabelas que não serão trabalhadas e por fim uma nova visualização na dataframe que recebe o novo valor.

```
dataFrame = rename(dataFrame, salario = salary_in_usd, cargo = job_title)
```

```
View(dataFrame)
```

esse comando de rename usei para ajudar nos acessos as tabelas, facilitando a compreensão dos títulos e o acesso aos mesmos.

Os demais comandos foram basicamente tentando acompanhar o que foi passado no workshop, porém mudei os nomes das variáveis para tentar compreender melhor os acessos as variáveis e os parâmetros, que no caso aqui é o dataframe e as colunas.

Na parte que começa a construção de classes eu parei um tempo também tentando entender os valores e de onde vinham, como o k, n, A, Amenor, A maior. Tive apenas uma apresentação de erro, pesquisei e parece ter a ver com as configurações da IDLE, resolvi parar por aqui.