

## **SCRIPT – DESAFIO R**

- Realizei o download e instalação do interpretador da linguagem R e a IDE RStudio e tomar como base inicial o script da do workshop;
- Escolhi um arquivo csv referente as cidades do mundo onde possuem a maior parte da população mundial;
- Neste conjunto de dados, temos dados populacionais históricos para cada país/território do mundo por diferentes parâmetros, como tamanho da área do país/território, nome do continente, nome da capital, densidade, taxa de crescimento populacional, classificação com base na população, Percentual da População Mundial etc.;
- O primeiro passo foi importar o arquivo csv para o RStudio, onde busquei entender melhor o comando `setwd`, usado definir o diretório de trabalho;
- Percebi que é sempre bom verificar a estrutura dos dados antes de usar as funções;
- Usando o comando `dim()` obtive exatamente a quantidade de linhas e colunas do data frame;
- Com o intuito de saber os nomes das colunas para posteriormente renomear as que eu considerar necessário, utilizei o comando `names()`;
- Tive dificuldades para conseguir usar o comando `rename()` mas erro foi solucionado após a instalação do pacote `dplyr`;
- Desde o desafio senti um pouco de receio, embora seja semelhante a python, nunca tive contato com R. Os primeiros erros começaram a aparecer, mas aos poucos, com determinação e seguindo as orientações que nos foram passadas, fui superando os erros e sanando as dúvidas uma a uma. Fico imensamente satisfeita por, aos poucos, aprender a usar o poder da informação, como também, desenvolver habilidades que serão fundamentais para minha vida profissional;
- A partir do pacote `dplyr` tentei utilizar o método `select()` para selecionar e remover colunas através do seu nome mas não deu certo. Então usei a palavra reservada `NULL`;
- Utilizando a função `mean()` pude construir um vetor que representa a média populacional mundial de 1970 até 2020. A partir desses dados irei realizar a estatística descritiva;
- Tive dificuldade de criar o data frame pois o console estava me retornando que a quantidade de argumentos estava diferente, então transformei os anos em strings e solucionei o problema. Mas não entendi o que aconteceu;
- Na realização da estatística descritiva, entendi as frequências absolutas e relativas, as demais não consegui fazer devido o prazo. Pretendo pesquisar mais e assim finalizar a estatística por completo.