

Desafio Opcional: Exploração do projeto alternativo

Nesta fase, seja curioso e tente descobrir o máximo que puder sobre os dados. Veja algumas sugestões:

- Conhecer a estrutura do *data frame* criado:

```
str(df_medidas)
```

- Obter diversas métricas sobre as variáveis:

```
str(df_medidas)
```

- Quantas observações temos por gênero e por movimento:

```
table(df_medidas$Genero)
table(df_medidas$Movimento)
```

- Procurar alguma correlação entre as coordenadas (valores negativos indicam correlação inversa, isto é, quando um cresce o outro diminui; positivos indicam crescimento uniforme, ou seja, ambas crescem ou ambas diminuem; os valores vão de -1 a +1 e quanto mais próximo do valor absoluto 1, mais forte a ligação entre as duas variáveis)

```
cor(df_medidas[c("X", "Y", "Z")])
```

Perceba que a correlação de uma variável com ela mesma sempre será 1.

- Monte um histograma revelando a distribuição das medidas obtidas nas coordenadas X:

```
hist(df_medidas$X, main = 'Histograma', xlab = 'X')
```

- Gere um gráfico com ocorrências por tipo de movimento:

```
library(sqldf)
Medidas_por_Tipo = sqldf("select Movimento, count(*) Qtos
                          from df_medidas group by Movimento
                          order by 2 desc")

library(ggplot2)
ggplot(Medidas_por_Tipo, aes(x = Movimento, y = Qtos)) +
  geom_col(aes(fill = Movimento))
```

Opinião do Instrutor

Continue com os seus estudos, e se houver dúvidas, não hesite em recorrer ao nosso fórum!