

**UNIVERSIDADE REGIONAL DE BLUMENAU**  
**Pós-Graduação em Data Science**  
**Disciplina Técnicas Estatísticas Aplicadas a Predição**  
**Prof. Moacir Manoel Rodrigues Junior**

**Exercício**

**Cenário**

Utilizando a planilha de dados disponibilizada calcula:

| <b>Rótulo</b> | <b>O que mede</b>                                    |
|---------------|--|
| obs           | Rótulo das Observações.                              |
| salario       | Salário dos indivíduos                               |
| idade         | Idade do Entrevistado                                |
| stress        | Nível de Stress do Entrevistado (escala de 0 a 100). |
| sexo          | 1 se mulher, 0 se homem                              |
| est_civil     | 1 se casado, 0 se solteiro                           |

**Questões:**

1. As medidas de tendência central das variáveis numéricas (mínimo, máximo, 1º Quartil, 3º Quartil, mediana e média)
2. As medidas de dispersão das variáveis numéricas (amplitude, desvio-padrão, variância, coeficiente de variação, assimetria e curtose)
3. Calcule a média e a mediana do salário para homens e mulheres separadamente.
4. Calcule o desvio-padrão e o coeficiente de variação para homens e mulheres separadamente e verifique se existe diferença nas variações.
5. Calcule a média e a mediana do salário para casado e solteiros separadamente.
6. Calcule a média e a mediana do nível de stress para homens e mulheres separadamente.
7. Calcule a média e a mediana do nível de stress para casado e solteiros separadamente.
8. Faça um gráfico de dispersão que compare o nível de stress com o salário do entrevistado.
9. Faça um gráfico de dispersão que compare o nível de stress com a idade do entrevistado.
10. Faça um gráfico box-plot das variáveis numéricas e verifique se existe a presença de outliers.
11. Considerando a média e o desvio-padrão para as mulheres, verifique qual a probabilidade de uma mulher receber acima do 3º quartil. E qual as chances de um homem? Assuma a distribuição normal para os dados.
12. Considerando a distribuição normal, com a média e o desvio-padrão dos casados, quais as chances do nível de stress de um destes respondentes ser menor do que 50 (metade da escala).