

EXERCÍCIO 2

Utilizando a base de dados “dados_salarios.csv” pede-se:

1) Retire uma amostra de 300 observações utilizando a sintaxe abaixo:

```
# Amostra
set.seed(02051975) # COLOQUE O SEU DIA MÊS E ANO DE NASCIMENTO NA SET.SEED
base1 = base[sample(nrow(base), 300),]
```

com a amostra retirada, faça a análise dos itens abaixo.

2) Para a variável “salario_USD” calcule a média, mediana, percentil5, percentil 25, percentil 75, percentil 95, mínimo e máximo. Tire conclusões sobre esses valores;

3) Calcule a média e a mediana da variável “salario_USD” para cada categoria (2020, 2021, 2022) da variável “ano”. Tire conclusões e indique se a mediana é mais apropriada do que a média na comparação.

4) Calcule a média e a mediana da variável “salario_USD” para cada categoria (EM, MI, SE, EX) da variável “experiencia”, porém antes da fazer a análise junte as categorias SE e EX, ou seja, analise a variável “experiencia” considerando 3 categorias. Tire conclusões e indique se a mediana é mais apropriada do que a média na comparação.

5) Apresente os resultados do item (4) em um gráfico utilizando os valores da mediana.

DESCRIÇÃO DA BASE DE DADOS

SALÁRIO DE FUNCIONÁRIOS QUE TRABALHAM COM DATA SCIENCE	
https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries	
Variável	Descrição
ano	O ano em que o salário foi pago.
experiencia	O nível de experiência no cargo durante o ano com os seguintes valores possíveis: EN: Nível básico MI: Nível médio SE: Nível sênior EX: Nível executivo
emprego	O tipo de emprego para a função: PT: Part-time FT: Full-time CT: Contract FL: Freelance
cargo	A função exercida durante o ano
salario_USD	O salário em USD (taxa de câmbio dividida pela taxa média em USD para o respectivo ano via fxdata.foorilla.com).
pais_empreg	O país de residência do funcionário durante o ano de trabalho como um código de país ISO 3166.
trab_remoto	O tempo total de trabalho feito remotamente, os valores possíveis são os seguintes: 0: Nenhum trabalho remoto (menos de 20%) 50: Parcialmente remoto 100: Totalmente remoto (mais de 80%)
pais_empresa	O país da sede do empregador ou da filial contratante como um código de país ISO 3166.
tam_empresa	O número médio de pessoas que trabalharam para a empresa durante o ano: S: menos de 50 funcionários (pequeno) M: 50 a 250 funcionários (médio) L: mais de 250 funcionários (grande)

Observação: Em um arquivo PDF, coloque a sintaxe do R, os resultados e as conclusões de cada tópico e poste no Moodle em EXERCICIO2 até às 8:00hs do dia 9/4. Não esqueça de colocar o nome de todos os membros do grupo que participaram da resolução do exercício.