

Questionário 1

Leia os dados com erro (iris-with-errors.csv), faça a limpeza e remova as duas últimas colunas. Quantas linhas e colunas sobraram nos dados?

- ☐ a. 12,3
- ☐ b. 14,3
- ☐ c. 14,4
- ☐ d. 14,10
- ☐ e. 12,2

Leia os dados com erro novamente e insira a media de cada atributo onde for encontrada NaN e '?'. Quais os valores das medianas obtidas para os atributos?

- ☐ a. [5.05 3.45 1.4 0.2]
- ☐ b. [5.12 4.45 12.4 2.2]
- ☐ c. [6.01 2.45 3.4 1.2]
- ☐ d. [4.15 3.45 2.4 1.2]
- ☐ e. [5.15 5.45 3.4 4.2]

Monte um boxplot para cada variável dos dados da Iris. Quais atributos apresentam outliers?

- ☐ a. petal_length
- ☐ b. sepal_width e petal_length
- ☐ c. sepal_width
- ☐ d. Todos.
- ☐ e. sepal_length e sepal_width

Considere a base de dados Advertising.csv. Calcule a correlação entre as variáveis. Quais são as variáveis mais correlacionadas?

- ☐ a. TV e radio
- ☐ b. radio e newspaper
- ☐ c. TV e newspaper
- ☐ d. radio e sales
- ☐ e. TV e sales