

Introdução à Ciência de Dados

Projeto 1

Enviar pelo Moodle (<https://edisciplinas.usp.br/course/view.php?id=82515>).

Considere o notebook desenvolvido na aula.

- 1 - Leia os dados com erro novamente, faça a limpeza e remova as duas últimas colunas.
- 2 - Leia os dados com erro novamente e insira a mediana de cada atributo onde for encontrada NaN.
- 3 - Considere os dados da Iris e mostre a distribuição de probabilidades de cada uma das variáveis após a normalização e padronização.
- 4 - Amostre os dados da Iris e selecione 10 elementos em cada classe.
- 5 - Monte um boxplot para cada variável dos dados da Iris.
- 6 - Encontre os outliers nos dados das bases Iris e BostonHouse.