



UNIVERSIDADE
CATÓLICA
PORTUGUESA

BRAGA

Machine Learning

Session 2 - PL

Data Preprocessing

Ciência de Dados Aplicada

2023/2024

Python libraries for data preprocessing

- **NumPy** is a fundamental Python library for numerical computing, offering powerful tools for efficient handling, transformation, and manipulation of multidimensional arrays;
- Documentation: www.numpy.org
- Tutorial: <https://numpy.org/doc/stable/user/quickstart.html>

Python libraries for data preprocessing

- **Pandas** is a versatile Python library for data manipulation and analysis, offering intuitive data structures along with extensive functionality for data preprocessing, exploration, and manipulation;
- Documentation: <https://pandas.pydata.org/docs/>
- Tutorial: https://pandas.pydata.org/docs/user_guide/10min.html

Exercises:

- Notebooks on the github repository:
 - Notebook with examples:
 - ❑ `notebooks/session2/examples.ipynb`
 - Notebook with exercises:
 - ❑ `Notebooks/session2/exercises.ipynb`