

Limpeza e transformação de dados em Python

Criada em 2020-10-29

DESIGNAÇÃO DA UFCD: Limpeza e transformação de dados em Python

CÓDIGO DA UFCD: 10808

CARGA HORÁRIA: 25 horas

PONTOS DE CRÉDITO: 2.25

Objetivos

- Proceder à importação e execução de queries às tabelas de dados.
- Executar operações básicas sobre os dados.
- Executar operações de limpeza de dados.
- Executar operações de transformação de dados e criação de variáveis.

Conteúdos

- Introdução às bibliotecas para análise de dados
 - Pandas
 - *NumPy*
 - *Scikit-Learn*
- Pandas
 - Data frames
 - Operações com *data frames (queries)*
 - Manipulação de diferentes estruturas e fontes de dados
- Ficheiros de texto
- *Comma-separated values*
- Folha de Cálculo
- Bases de dados
 - Construção de variáveis

- *NumPy*
 - Matrizes e vetores
 - Operações básicas com matrizes e vetores
 - Produção de estatísticas descritivas
- *Scikit-Learn*
 - Normalização de dados
 - Identificação e imputação de valores omissos
 - Identificação e remoção de valores extremos
 - Variáveis categóricas
 - Redução de dimensionalidade
 - Exploração de dados
- Projeto de preparação e pré-processamento de dados