

Identificação da Disciplina

Professor(a) responsável: Ramon da Silva Torres

Nome do Curso: Introdução ao Python para Ciência de Dados

Carga horária: 30 horas

Nível: Pós-Graduação

Pré-requisito: Inglês Básico

Objetivos

Objetivo geral: Capacitar os alunos às operações com bases de dados através de códigos na linguagem de programação Python e da interface do Google Colab.

Objetivos específicos:

- Operar dados tabulares com python, realizando limpeza nos dados e transformações;
- Construir visualizações e rodar modelos estatísticos.

Conteúdo Programático

- Python e as demais ferramentas de bases de dados;
- Carregando dados tabulares e obtendo informações da base;
- Operando transformações nos dados;
- Visualizações, testes estatísticos e regressões lineares;
- Construindo Pipeline de dados.
- Automatização

Metodologia

Aulas práticas: Serão desenvolvidas por meio de exposição oral com demonstração dos códigos e boas práticas de programação em conjunto com exercícios a serem realizados pelos alunos. As aulas serão compostas blocos de 1:30h cada (dois por turno). Serão disponibilizados aos alunos material de apoio para o curso em slides e em código.

Bibliografia

- MCKINNEY, W. Python for Data Analysis: Data Wrangling with Pandas, Numpy, and IPython. O'Reilly, 2017.
- GRUS, J. Data Science from Scratch: first principles with python. O'Reilly, 2015
- VANDERPLAS, J. Python Data Science Handbook. O'Reilly, 2017.
- RAMALHO, L. Fluent Python: Clear, Concise, and Effective Programming. O' Reilly, 2nd Edition, 2022.
- SCAVETTA, R & ANGELOV, R. Python and R for the Modern Data Scientist: The best of both Worlds. O'Reilly, 2021.