



Lisbon School
of Economics
& Management
Universidade de Lisboa



Academic Year: 2022/2023

PROGRAMMING FOR DATA SCIENCE

Apresentação

- Professor
- Learning Objectives
- Program
- Bibliography
- Evaluation rules

Instructor



Prof. Carlos J. Costa, PhD

email: cjcosta@iseg.ulisboa.pt

<https://www.iseg.ulisboa.pt/aquila/homepage/cjcosta>

<https://scholar.google.com/citations?user=CpxlHn0AAAAJ&hl=en>

<https://www.youtube.com/channel/UC24JG7YrzZs1xMNFkp73xuw/videos>

slideshare: carlosjcosta

GitHub: masterfloss

Twitter: @carlosjcosta

Instagram: kjkosta

Office:

Prof. Carlos J. Costa

Rua Miguel Lupi, nº 20 – gabinete 318

Tuesday – 17:00 to 19:00

Course Objectives

The student should obtain the following skills:

- Consolidate main programming concepts
- Understand programming techniques to manipulate and visualize data
- Use programming languages to explore data
- Create models supported in data.

Course Syllabus

I. Overview to Programming Concepts

II. Main libraries used in Data Science.

1. data analysis and statistics

2. data visualization

3. Networks

4. image manipulation and text

5. machine learning

6. Deployment and Web

III. Data Science.

Projects Presentation

Projects Presentation

Bibliography

- Albon, C. (2018). *Machine learning with python cookbook: Practical solutions from preprocessing to deep learning*. " O'Reilly Media, Inc."
- Bird,S, Klein, E. & Loper, E. (2019) Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit <http://www.nltk.org/book/>
- Hagberg, A., Swart, P. J., & Schult, D. A. (2020) NetworkX Reference, Release 2.5 https://networkx.org/documentation/stable/_downloads/networkx_reference.pdf
- Martins J. P. (2015) Programação em PYTHON: Introdução à Programação Utilizando Múltiplos Paradigmas, IST Press.
- McKinney, W., & Team, P. D. (2021). Pandas-powerful Python data analysis toolkit 1.2.2. Download from: <https://pandas.pydata.org/docs/pandas.pdf>

Libraries websites:

1. <https://numpy.org/>
2. <https://pandas.pydata.org/>
3. <https://www.statsmodels.org/stable/index.html>
4. <https://scikit-learn.org/stable/>
5. <https://matplotlib.org/stable/index.html>
6. <https://seaborn.pydata.org/>
7. <https://networkx.org/>
8. <http://www.nltk.org/>
9. <https://scikit-image.org/>
10. <https://opencv.org/>

Evaluation

- Test: 30%
- Quiz and labs: 30%
- Team Work: 40%

Deadlines

1. Team registration (9/03)
2. Project Statement (21/03)
3. Project Proposal presentation (30/03)
3. Project Submission (11/05)
5. Project Presentation (16 and 18/05)

What I expect from students

- Classes must be complemented with individual study hours.
- Act with ethics towards work and others.
- Learn and participate in class activities with enthusiasm
- Be respectful
 - *listen | share airtime | open mind | use of personal devices only when required or during breaks*
- Be Responsible:
 - *Arrive on time | follow class activities | help others (but do not do their work) | integrate your colleagues in the group work*
- Be a problem solver
 - *ask questions | Share ideas | embrace the struggle of learning | Stay positive!*

AI Support

- You have many tools that may be used (Chat-GPT, Chatsonic, You.com, jasper.ai,)
- But must be used ethically