COURSE OUTLINE

GENERAL

SCHOOL	Sciences and Engineering		
ACADEMIC UNIT	Computer Science		
LEVEL OF STUDIES	1st Cycle		
COURSE CODE	COMP-240	**SEMESTER**	Fall
COURSE TITLE	Data Programming		

INDEPENDENT TEACHING ACTIVITIES *if credits are awarded for separate components of the course, e.g. lectures, laboratory exercises, etc. If the credits are awarded for the whole of the course, give the weekly teaching hours and the total credits*	**WEEKLY TEACHING HOURS**	**CREDITS**
	3.5	6
Add rows if necessary. The organisation of teaching and the teaching methods used are described in detail at (d).		

COURSE TYPE *general background, special background, specialised general knowledge, skills development*	special background
PREREQUISITE COURSES:	COMP-111
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	English
IS THE COURSE OFFERED TO ERASMUS STUDENTS	
COURSE WEBSITE (URL)	

LEARNING OUTCOMES

Learning outcomes

The course learning outcomes, specific knowledge, skills and competences of an appropriate level, which the students will acquire with the successful completion of the course are described.

Consult Appendix A

- *Description of the level of learning outcomes for each qualifications cycle, according to the Qualifications Framework of the European Higher Education Area*
- *Descriptors for Levels 6, 7 & 8 of the European Qualifications Framework for Lifelong Learning and Appendix B*
- *Guidelines for writing Learning Outcomes*

After completion of the course students are expected to be able to:

- Understand the basic concepts involved in computational thinking towards solving data-oriented problems.
- Distinguish between suitable data types and structuring techniques for different data programming challenges.
- Develop programs in Python for data processing.
- Distinguish between suitable programming paradigms and evaluate strengths and weaknesses of each paradigm depending on the context of different Data Science tasks.
- Apply in practice the principles for pre-processing datasets and for indexing data to improve

the efficiency of data trawling and extraction.

- Understand the algorithmic process for cleaning impure datasets and acknowledge the scientific treatments for dealing with missing data.
- Apply in practice the principles for combining, aggregating, and summarizing datasets.
- Acknowledge how the data collection process influences the scope of inference.
- Comprehend basic methods towards plotting data in order to present insights extracted from datasets.
- Be familiar with the methods for measuring and optimizing the performance of data-intensive tasks to cope with large volumes of data

General Competences

Taking into consideration the general competences that the degree-holder must acquire (as these appear in the Diploma Supplement and appear below), at which of the following does the course aim?

Search for, analysis and synthesis of data and information, with the use of the necessary technology	*Project planning and management*
Adapting to new situations	*Respect for difference and multiculturalism*
Decision-making	*Respect for the natural environment*
Working independently	*Showing social, professional and ethical responsibility and sensitivity to gender issues*
Team work	*Criticism and self-criticism*
Working in an international environment	*Production of free, creative and inductive thinking*
Working in an interdisciplinary environment	*……*
Production of new research ideas	*Others…*
	…….

Search for, analysis and synthesis of data and information, with the use of the necessary technology
Adapting to new situations
Decision-making
Working independently
Team work
Production of free, creative and inductive thinking

SYLLABUS

1. Programming for Data Science with Python
    1. Iterative Programming
    2. Functional Decomposition
    3. Error Handling
    4. Data-Oriented Problem Solving
2. Associative Data Types and Structures
    1. Lists
    2. Dictionaries
    3. Tuples
    4. Sets
3. Interactive Programming Environments
    1. Jupyter Notebooks

2. Markup
3. Version Control

4. Arrays and Vectorized Computation
    1. Multi-dimensional Array Objects (e.g., NumPy)
    2. Array-Oriented Programming
    3. Vectorization for Arithmetic Operations
    4. Array Broadcasting
5. Tabular and Heterogeneous Data
    1. Tabular Objects (e.g., DataFrames)
    2. Hierarchical Data Indexing
    3. Data Grouping, Aggregation and Summarization
6. Classes and Objects
    1. Data Abstraction and Encapsulation
    2. Inheritance
    3. Polymorphism
7. Dealing with Data Sources
    1. Strings and Encoding Schemes
    2. Regular Expressions
    3. File I/O
    4. Web Scraping
8. Data Pre-Processing
    1. Data Cleaning
    2. Lambda Transformations and Functional Programming
    3. Attribute Scaling
    4. Categorical Data Encoding
9. Data Wrangling
    1. Data Reshaping and Pivoting
    2. Sampling and Filtering
    3. Data Grouping and Aggregation
    4. Compression for Tabular Data
10. Statistical Modeling
    1. Descriptive Statistics
    2. Simulation-Based Experimentation
    3. Discovering Data Correlations

TEACHING and LEARNING METHODS - EVALUATION

DELIVERY *Face-to-face, Distance learning, etc.*	Face-to-face
USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY *Use of ICT in teaching, laboratory education, communication with students*	*Use of ICT in teaching / Χρήση ΤΠΕ* *Communication with students / Επικοινωνία με Φοιτητές*
TEACHING METHODS *The manner and methods of teaching are described in detail.* *Lectures, seminars, laboratory practice, fieldwork, study and analysis of bibliography, tutorials, placements, clinical practice, art workshop, interactive teaching, educational visits, project, essay writing, artistic creativity, etc.* *The student's study hours for each learning activity are given as well as the hours of non-directed study according to the principles of the ECTS*	
STUDENT PERFORMANCE EVALUATION *Description of the evaluation procedure* *Language of evaluation, methods of evaluation, summative or conclusive, multiple choice questionnaires, short-answer questions, open-ended questions, problem solving, written work, essay/report, oral examination, public presentation, laboratory work, clinical examination of patient, art interpretation, other* *Specifically-defined evaluation criteria are given, and if and where they are accessible to students.*	Homework Assignments, Midterm Exam, Participation, Final Exam, Data Programming Challenge

Activity	***Semester workload***
Lectures	49
Preparation	24
Coursework	30
Exam Preparation	43
Examination	4
Course total	***150***

ATTACHED BIBLIOGRAPHY

Required Textbooks / Readings:

Title	**Author(s)**	**Publisher**	**Year**	**ISBN**
Python for Data Analysis* (3rd edition)	Wes McKinney	O'Reilly Media	2022	978-1-098-10403-0
Think Python** (2nd edition)	Allen B. Downey	Green Tree Press	2016	978-1-491-93936-9

* Also available as an open access e-book by the author https://wesmckinney.com/book/

** A free pdf version is available by the author https://greenteapress.com/wp/think-python-2e/

Recommended Textbooks / Readings:

Title	**Author(s)**	**Publisher**	**Year**	**ISBN**

Computational and Inferential Thinking: The Foundations of Data Science (2nd edition)	Ani Adhikari, John DeNero and David Wagner	OpenLibra	2022	9124423963
Python Data Science Handbook	Jake VanderPlas	O'Reilly Media	2016	978-1-491-91205-8
Doing Data Science	Cathy O'Neil and Rachel Schutt	O'Reilly Media	2014	978-1-449-35865-5