UNIVERSITY of NICOSIA

Course Syllabus

Course Code	Course Title	ECTS Credits
COMP-542DL	Data Programming	10
Prerequisites	**Department**	**Semester**
None	Computer Science	Fall
Type of Course	**Field**	**Language of Instruction**
Required	Data Science	English
Level of Course	**Lecturer(s)**	**Year of Study**
2nd Cycle	Dr. Ioannis Partalas	1st
Mode of Delivery	**Work Placement**	**Corequisites**
Distance Learning	N/A	None

Course Objectives:

The main objectives of the course are to:

- Present fundamental programming concepts, data abstractions and computing techniques for Data Science.
- Introduce a problem-solving framework for "thinking with data".
- Present intuitive data types and effective data structuring techniques to speed data curation.
- Introduce programming paradigms suitable for Data Science projects and in which contexts are these paradigms embraced.
- Provide the essential programming skills for Data Science using Python.
- Introduce the basics of the Python programming language.
- Introduce the different tools for data harvesting and manipulation.

Learning Outcomes:

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

1. Explain the basic concepts involved in computational thinking towards solving data-oriented problems.
2. Distinguish between suitable data types and structuring techniques for different data programming challenges.
3. Develop relatively well-structured and maintainable programs in Python.
4. Distinguish between suitable programming paradigms and evaluate strengths and

weaknesses of each paradigm depending on the context of different Data Science tasks.
5. Apply in practice algorithmic methods for data harvesting and curation.
6. Use different core libraries of Python.

Course Content:

1. Introduction to Data Programming
    a. The basic Concepts of Data Programming
    b. Data Programming Lifecycle
    c. The toolsets of a Data Scientist
2. Introduction to Data Programming II
    a. Interacting with Python
    b. Python Syntax
    c. Variables Data Types
    d. Control flow
    e. Conditional expressions
3. Introduction to Data Programming III
    a. Collections in Python
    b. Linear data Structures
    c. Associative Data Structures
4. Classes and Objects
    a. Classes and Objects
    b. Inheritance
    c. Encapsulation
    d. Data Abstractions
5. Interactive Programming
    a. Interactive Programming
    b. Creating, Publishing and Sharing Notebooks
6. Introduction to Programming with Arrays and Matrices I
    a. Array-Oriented Programming
    b. Numerical Arrays
    c. Indexes and Arrays
    d. Multi-Dimensional Arrays
    e. Reshaping, Joining and Splitting Arrays
7. Introduction to Programming with Arrays and Matrices II
    a. Advanced Array Operators
    b. Non-Homogeneous Arrays
    c. Sorting Arrays
8. Handling Data Sources
    a. Strings and Encoding Schemes
    b. Regular Expressions
    c. File I/O
    d. The CSV and JSON formats
    e. Manipulating Textual Data
9. Data Manipulation and DataFrames I
    a. Dataframe Objects

UNIC

b. Hierarchical Indexes
c. Data Grouping, Aggregation and Summarization

10. Data Manipulation and DataFrames II
    a. Missing Values in DataFrames
    b. DataFrame Filling Strategies
11. Functional Programming
12. From Data Programming to Data Mining

Learning Activities and Teaching Methods:

Lectures, Exercises, Software Tool Tutorials, Case-Study Presentations, Discussions.

Assessment Methods:

Homework, Projects, Final Assessment*

* The Final Assessment can be either a Final Exam or Final Assignment(s) with Viva

Required Textbooks / Readings:

Python Data Science Handbook	Jake VanderPlas	O'Reilly Media	2016	978-1-491-91205-8

Recommended Textbooks / Readings:

Think Python* (2nd edition)	Allen B. Downey	Green Tree Press	2015	978-1-491-93936-9
Think Like a Data Scientist	Brian Godsey	Manning	2017	978-1-633-43027-3