COURSE OUTLINE

GENERAL

SCHOOL	Sciences and Engineering		
ACADEMIC UNIT	Computer Science		
LEVEL OF STUDIES	1st Cycle		
COURSE CODE	COMP-140	**SEMESTER**	Fall
COURSE TITLE	Introduction to Data Science		

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**
	2.5	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:	None
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:

- Describe the importance and value of Data Science to society, science and business.
- Acknowledge the different roles of a Data Scientist and the skills required for each role.
- Distinguish between the different types of data and identify basic challenges related to "big" data.
- Describe the stages, tasks and problems involved in a Data Science project.
- Understand the basic concepts involved in statistical thinking towards solving data- oriented problems.
- Recognize the different toolsets required to attack different Data Science projects.

- Comprehend the methods towards presenting stories not immediately evident within datasets and knowledge extracted from raw data.
- Describe the Machine Learning motif and be able to distinguish among different Data Mining techniques (i.e., regression, classification, clustering) in relevance to a given problem description.
- Identify and understand the ethical and privacy concerns for society when processing sensitive and personal 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
Working in an interdisciplinary environment
Production of free, creative and inductive thinking

SYLLABUS

1.The Multi-Faceted Profile of a Data Scientist
   a. Digital Systems and Datafication
   b. The Current Data Science Landscape and Future Directions
   c. The Different Roles and Skillsets of a Data Scientist
   d. Data-Driven Decision-Making
2.Data and Datasets
   a. The Origins of Data
   b. Attribute Types and Values
   c. Dataset Structure-wise Decomposition
   d. Big Data
3.The Data Science Process
   a. From Raw Data to Analytics Insights and Knowledge
   b. The Different Stages of the Data Science Process
   c. The Toolsets of a Data Scientist
4.Statistical and Algorithmic Thinking
   a. Data-Oriented Problem Solving
   b. Descriptive Statistics

c. The Philosophy of Exploratory and Explanatory Data Analysis

5.Data Visualization
   a. Basic Dataset Plotting
   b. Data Storytelling
   c. Data Visualization as a Data Science Tool

6.Observation Studies
   a. Correlation and Causality
   b. Confounding Factors
   c. Sampling in the Data Collection Process

7.Data Pre-Processing
   a. Missing Data
   b. Outliers
   c. Record Similarity

8.Data Mining
   a. The Importance of Models in Data Science
   b. Algorithmic Modeling and Inference
   c. Data Mining Techniques
   d. The Machine Learning Motif

9.The Ethical Concerns Faced by Data Scientists
   a. Ethics in the Digital Era
   b. Ethical and Privacy Debates

10.Data Science Case Studies

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*	

Activity	***Semester workload***
Lectures	35
Preparation	26
Coursework	40
Exam Preparation	45
Examination	4
Course total	***150***

STUDENT PERFORMANCE EVALUATION *Description of the evaluation procedure*	Homework Assignments, Midterm Exam, Project, Participation, Final Exam

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.*	

ATTACHED BIBLIOGRAPHY

Required 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
Doing Data Science	Cathy O'Neil and Rachel Schutt	O'Reilly	2014	978-1-449-35865-5

***** Also, available as an open access e-book by the authors at https://inferentialthinking.com/

Recommended Textbooks / Readings:

Title	Author(s)	Publisher	Year	ISBN
Think Like a Data Scientist	Brian Godsey	Manning	2017	978-1-633-43027-3
Data Science	John D. Kelleher and Brendan Tierney	MIT Press	2018	978-1-469-09677-3
The Art of Data Science	Roger D. Peng and Elizabeth Matsui	LuLu	2016	978-1-365-06146-2