

ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

103000888 - Programming For Data Science

DEGREE PROGRAMME

10AZ - Master Universitario En Innovación Digital

ACADEMIC YEAR & SEMESTER

2023/24 - Semester 2

Index

Learning guide

1. Description.....	1
2. Faculty.....	1
3. Skills and learning outcomes	2
4. Brief description of the subject and syllabus.....	3
5. Schedule.....	4
6. Activities and assessment criteria.....	6
7. Teaching resources.....	7
8. Other information.....	8

1. Description

1.1. Subject details

Name of the subject	103000888 - Programming For Data Science
No of credits	5 ECTS
Type	Optional
Academic year of the programme	First year
Semester of tuition	Semester 2
Tuition period	February-June
Tuition languages	English
Degree programme	10AZ - Master Universitario en Innovación Digital
Centre	10 - Escuela Técnica Superior De Ingenieros Informáticos
Academic year	2023-24

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Guillermo Antonio Viguera Gonzalez (Subject coordinator)	D-4310	guillermo.viguera@upm.es	Tu - 10:00 - 13:00 Th - 10:00 - 13:00
Raul Gutierrez Gil	D4310	r.gutierrez@upm.es	M - 15:00 - 18:00 W - 09:00 - 12:00

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

2.3. External faculty

Name and surname	Email	Institution
Belén Ríos Sánchez	belen.rios@upm.es	Escuela Técnica Superior de Ingenieros Informáticos

3. Skills and learning outcomes *

3.1. Skills to be learned

CB06 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB07 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CE-CD05 - Capacidad para usar herramientas de procesamiento de big data tanto en online como en modo batch

CE-CD06 - Capacidad para extraer, integrar y consultar datos heterogéneos en diferentes escenarios

3.2. Learning outcomes

RA54 - Poseer destrezas fundamentales de la programación que permitan la implementación de algoritmos y el uso de estructuras de datos típicos en ciencia de datos. e distintos tipos de herramientas (software o metodológicas y conceptuales) necesarias para el correcto y eficaz desarrollo de software, incluyendo entornos y librerías en el contexto de ciencia de datos.

RA55 - Conocimiento y aplicación de algoritmos y estructuras de datos básico, así como las técnicas y métodos generales para su diseño.

RA53 - Destreza en el uso de distintos tipos de herramientas (software o metodológicas y conceptuales) necesarias para el correcto y eficaz desarrollo de software, incluyendo entornos y librerías en el contexto de ciencia de datos.

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

4. Brief description of the subject and syllabus

4.1. Brief description of the subject

This course is related with data manipulation and programming using the Python language. The main goal is to introduce main characteristics and programming style using this language of wide adoption for data analysis purposes. Once introduced the language, the course presents to students how to efficiently use the different constructs, control statements and data structures in Python. Based on this, the course describes the main characteristics of a Python framework for data storage and manipulation. All programming concepts presented in the course are accompanied with exercises in order to guarantee correct comprehension and practical knowledge.

4.2. Syllabus

1. Python programming.
 - 1.1. Introduction to Python programming.
 - 1.2. Common Python data structures.
 - 1.3. Programming in Python adopting an imperative paradigm.
 - 1.4. Programming in Python adopting a functional paradigm.
2. Data representation and manipulation.
 - 2.1. Python framework for data representation.
 - 2.2. Revision of data manipulation techniques.
3. Data Visualization.
4. Scalable processing of large scale datasets
 - 4.1. Introduction to large scale data processing
 - 4.2. Framework for large scale data processing
 - 4.3. Programming techniques for large scale data processing

5. Schedule

5.1. Subject schedule*

Week	Classroom activities	Laboratory activities	Distant / On-line	Assessment activities
1	Python programming Duration: 02:00			
2	Python programming Duration: 02:00	Python programming Duration: 02:00		
3	Python programming Duration: 02:00	Python programming Duration: 02:00		
4	Python programming Duration: 02:00	Python programming Duration: 01:00		
5	Data Manipulation Duration: 02:00	Data Manipulation Duration: 02:00		
6	Data Manipulation Duration: 02:00	Data Visualization Duration: 02:00		
7	Scalable data processing Duration: 02:00	Scalable data processing Duration: 02:00		
8	Scalable data processing Duration: 02:00	Scalable data processing Duration: 01:00		Submit Programming Exercise Continuous assessment Not Presential Duration: 01:00 Submit Programming Exercise Final examination Not Presential Duration: 01:00
9				
10				
11				
12				
13				
14				
15				
16				

17				
----	--	--	--	--

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

* The schedule is based on an a priori planning of the subject; it might be modified during the academic year, especially considering the COVID19 evolution.

6. Activities and assessment criteria

6.1. Assessment activities

6.1.1. Assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
8	Submit Programming Exercise		No Presential	01:00	100%	5 / 10	CB06 CB07 CE-CD05 CE-CD06

6.1.2. Global examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
8	Submit Programming Exercise		No Presential	01:00	100%	5 / 10	CB06 CB07 CE-CD05 CE-CD06

6.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Submit Programming Exercise		Face-to-face	01:00	100%	5 / 10	CB06 CB07 CE-CD05 CE-CD06

6.2. Assessment criteria

Sistema general de evaluación

Evaluation system during the course will consist of developing one practical exercise performed during the course. The weight in the final grade of the exercise is 100% and a minimum of 5 points out of 10 is required in order to pass the course.

Sistema de Evaluación Global

Global evaluation system will consist of developing one practical exercise to be submitted on the day scheduled for the course evaluation during for the global evaluation system. The weight in the final grade of the exam is 100% and a minimum of 5 points out of 10 is required in order to pass the course.

Sistema de Evaluación Convocatoria extraordinaria

Evaluation system during the extraordinary period will consist of developing one practical exercise to be submitted on the day scheduled for the extraordinary period. The weight in the final grade of the exam is 100% and a minimum of 5 points out of 10 is required in order to pass the course.

7. Teaching resources

7.1. Teaching resources for the subject

Name	Type	Notes
Moodle	Web resource	Main communication channel with students. Repository: slides, scripts, data sets and other resources

Data Science from Scratch: First Principles with Python	Bibliography	Joel Grus
Programming in Python 3: A Complete Introduction to the Python Language	Bibliography	Mark Summerfield

8. Other information

8.1. Other information about the subject