Syllabus

Course Description

Course Title	Programming and Visualisation for Data Science
Course Code	27500
Course Title Additional	
Scientific-Disciplinary Sector	INFO-01/A
Language	English
Degree Course	Master in Data Analytics for Economics and Management
Other Degree Courses (Loaned)	Loaned from course 73081 - Master in Computing for Data Science (LM-18)
Lecturers	Prof. Antonio Liotta, Antonio.Liotta@unibz.it https://www.unibz.it/en/faculties/engineering/academic-staff/person/41903
Teaching Assistant	
Semester	First semester
Course Year/s	1
CP	12
Teaching Hours	M1: 40 hours M2: - 24 hours of in-person lectures - 12 hours of video lectures (counted as 24 hours to account for re-watching)
Lab Hours	40 (20 + 20)
Individual Study Hours	-
Planned Office Hours	M1: 18 hours M2: 18 hours
Contents Summary	Module 1 provides a comprehensive introduction to Python programming, starting with the setup of the development environment and core programming constructs. Students will explore Python's data structures and programming primitives, progressing to object-oriented programming and the development

	of structured, reusable code using functions, classes, and libraries. Emphasis is placed on best practices in software development, including code documentation, testing, version control, and distribution. The module concludes with advanced Python programming techniques, preparing students to build robust and maintainable applications. Module 2 guides students through the complete data science pipeline, from raw data acquisition to advanced analytics and visualization. Students will gain hands-on experience in data ingestion, exploration, cleaning, and feature engineering, building a strong foundation for effective data modeling. The course covers key machine learning techniques—including clustering, classification, and regression—alongside model tuning, validation, and testing. Emphasis is placed on producing insightful and reproducible visualizations using specialized Python libraries, enabling students to communicate data-driven findings with clarity and impact.
Course Topics	Module 1 is designed to provide specific professional skills for advanced programming in Python. The students will learn how to develop a Python program, starting from designing it, and going through coding, testing and validation. They will master Python in its full object-oriented features, learning how to develop complex programs that are well structured, and make use of techniques for code re-use, pipelining, maintenance, and deployment. Module 2 is designed to acquire professional skills and knowledge useful when dealing with large-scale datasets. In particular, the students will master data collection, exploration, transformation, curation, analysis, and visualization, choosing the most appropriate technique for the data at hand. They will make insights from the data, supported by a rigorous data science pipeline, which starts with raw data, produces machine learning models, and ends with advanced visualizations. This module, addresses common pitfalls that can mislead the analysis and makes extensive use of specialized Python libraries, acquiring the best practices of reproducible, data-driven analysis and research.
Keywords	Python Programming; data science; data curation; data modeling; machine learning; data visualization.

Recommended Prerequisites	Basic programming concepts. Linear algebra. Basic Statistics.
Propaedeutic Courses	
Teaching Format	The course adopts a blended, student-centered approach that emphasizes problem-based learning and active engagement. Selected lecture content is made available online in advance, enabling students to explore key concepts independently and at their own pace. This preparatory work allows in-person sessions to focus on applying knowledge through problem-solving, collaborative activities, and guided discussions—fostering critical thinking and deeper understanding. The teaching format combines frontal lectures, hands-on lab assignments, and project work, ensuring that students develop both theoretical knowledge and practical skills in python programming, and in data analysis, modelling, and visualization. The course is aligned with the principles of the EDUNEXT initiative (https://edunext.eu), promoted by Italian universities, which supports the integration of digital resources and active learning strategies in higher education.
Mandatory Attendance	Not compulsory. Non attending students have to agree with the lecturer on the modalities of independent study at the beginning of the course.
Specific Educational Objectives and Learning Outcomes	Intended Learning Outcomes (ILO) M1: ILO 1 Knowledge and understanding: ILO 1.1 The student acquires programming knowledge, particularly aimed at data analysis and statistical methodologies for implementing models as well as analysing large-scale datasets. In particular, the computing skills are focused on machine learning methods, on understanding modern techniques for data management and storage, including data from heterogeneous sources in terms of type and structure, such as spatio-temporal data and high-dimensional data, also in cloud environments, and on implementing algorithms for massive data processing. ILO 2 Applying knowledge and understanding: ILO 2.1 Students will develop the ability to apply and implement techniques

	for analysing large-scale datasets and spatio-temporal data under conditions of uncertainty, through the design and development of algorithms. The goal is to ensure the utility, quality, and effectiveness of the analysis. ILO 2.2 Ability to use IT technologies, techniques and methodologies for the acquisition, management, integration, analysis and visualisation of large datasets, in order to ensure scalability in terms of dataset volume and acquisition speed. These skills relate in particular to large database and dataset management systems and related visualisation techniques, models and languages for expressing data semantics, learning techniques, decision-making models, information systems organisation, web search techniques and data flow management techniques. ILO 3 Making judgements: ILO 3.1 The student acquires the ability to apply acquired knowledge to interpret data in order to make directional and operational decisions in a business context. ILO4 Communication skills: ILO 4.1 The student acquires the ability to communicate effectively in oral and written form the specialised content of the individual disciplines, using different registers, depending on the recipients and the communicative and didactic purposes, and to evaluate the formative effects of his/her communication. ILO 5 Learning skills: ILO 5.1 The student acquires knowledge of scientific research tools. He/she will also be able to make autonomous use of information technology to carry out bibliographic research and investigations both for his/her own training and for further education. Furthermore, through the curricular teaching and the activities related to the preparation of the final thesis, she will be able to acquire the ability - to identify thematic connections and to establish relationships between methods of analysis and application contexts; - to frame a new problem in a systematic manner and to implement appropriate analysis solutions; - to formulate general statistical-econometric models from the phenomena studied

	M2: ILO 1 Knowledge and understanding: ILO 1.1 The student acquires programming knowledge, particularly aimed at data analysis and statistical methodologies for implementing models as well as analysing large-scale datasets. In particular, the computing skills are focused on machine learning methods, on understanding modern techniques for data management and storage, including data from heterogeneous sources in terms of type and structure, such as spatio-temporal data and high-dimensional data, also in cloud environments, and on implementing algorithms for massive data processing. ILO 1.2 Students will acquire knowledge and skills in the analysis of textual data and network structures, with particular attention to issues related to data security and privacy. ILO 2 Applying knowledge and understanding: ILO 2.1 Students will develop the ability to apply and implement techniques for analysing large-scale datasets and spatio-temporal data under conditions of uncertainty, through the design and development of algorithms. The goal is to ensure the utility, quality, and effectiveness of the analysis. ILO 2.2 Ability to use IT technologies, techniques and methodologies for the acquisition, management, integration, analysis and visualisation of large datasets, in order to ensure scalability in terms of dataset volume and acquisition speed. These skills relate in particular to large database and dataset management systems and related visualisation techniques, models and languages for expressing data semantics, learning techniques, decision-making models, information systems organisation, web search techniques and data flow management techniques. ILO 3 Making judgements: ILO 3.1 The student acquires the ability to apply acquired knowledge to interpret data in order to make directional and operational decisions in a business context. ILO 3.2 The student acquires the ability to apply acquired

	knowledge to support processes related to production, management and risk promotion activities and investment choices through the organisation, analysis and interpretation of complex databases. ILO 4.1 The student acquires the ability to communicate effectively in oral and written form the specialised content of the individual disciplines, using different registers, depending on the recipients and the communicative and didactic purposes, and to evaluate the formative effects of his/her communication. ILO 5 Learning skills: ILO 5.1 The student acquires knowledge of scientific research tools. He/she will also be able to make autonomous use of information technology to carry out bibliographic research and investigations both for his/her own training and for further education. Furthermore, through the curricular teaching and the activities related to the preparation of the final thesis, she will be able to acquire the ability - to identify thematic connections and to establish relationships between methods of analysis and application contexts; - to frame a new problem in a systematic manner and to implement appropriate analysis solutions; - to formulate general statistical-econometric models from the phenomena studied.
Specific Educational Objectives and Learning Outcomes (additional info.)	
Assessment	The exam modalities are the same for both the attending and the non-attending students. All project works must have been submitted, at the very latest, 15 days ahead of the oral exam. In case of a positive mark, the projects will count for all 3 regular exam sessions. Module 1: Project work (70% of the final grade) (ILOs 1, 2, 5). Oral exam (30% of the final grade) (ILOs 3, 4). Module 2: Project work (70% of the final grade) (ILOs 1, 2, 5). Oral exam (30% of the final grade) (ILOs 3, 4).

Evaluation Criteria	70% project work, 30% oral exam. • Relevant for project work: clarity of presentation, ability to gain useful and novel insights from data, creativity, critical thinking, ability to adhere to reproducible research best practices • Ability to use Python to write, evaluate and deploy advanced, object-oriented computer programs • Ability to use Python to employ (understand, recall and use) data analytics methods in practical settings, from data collection and curation, to data analysis, modelling and visualization.
Required Readings	*Data Visualization. A practical introduction.* Haley. Available online *A layered grammar of graphics.* Wickham. Available online *Python Data Science Handbook*, by Jake VanderPlas. O'Reilly Media (1st Edition, 2016). Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it
Supplementary Readings	*Fundamentals of Data Visualization*. Wilke. Available online *Visualization Analysis and Design*. Munzer. Amazon *Data Visualization: Charts, Maps, and Interactive Graphics.* Grant. Amazon *Doing Data Science*. Cathy O'Neil, Rachel Schutt. O'Reilly, 2013, https://www.oreilly.com/library/view/doing-data-science/9781449363871/ *Python for Data Analysis.* By Wes McKinney. O'Reilly, 2nd Edition, 2017, https://www.oreilly.com/library/view/python-for-data/9781491957653/
Further Information	

Sustainable Development Goals (SDGs)	Industry, innovation and infrastructure, Quality education

Course Module

Course Constituent Title	M1 - Introduction to programming for data science
Course Code	27500A
Scientific-Disciplinary Sector	INFO-01/A
Language	English
Lecturers	
Teaching Assistant	
Semester	First semester
CP	6
Responsible Lecturer	
Teaching Hours	40
Lab Hours	20
Individual Study Hours	-
Planned Office Hours	18
Contents Summary	This module provides a comprehensive introduction to Python programming, starting with the setup of the development environment and core programming constructs. Students will explore Python's data structures and programming primitives, progressing to object-oriented programming and the development of structured, reusable code using functions, classes, and libraries. Emphasis is placed on best practices in software development, including code documentation, testing, version control, and distribution. The module concludes with advanced Python programming techniques, preparing students to build robust and maintainable applications.
Course Topics	This course provides students with advanced professional skills for developing robust and maintainable Python applications. It covers the full development lifecycle—from program design to implementation, testing, and validation. Students will master Python's object-oriented programming features and learn how to build well-structured, reusable, and modular code using functions, classes, and libraries. The course also introduces essential tools and techniques for code

	documentation, testing, version control, and distribution, preparing students for collaborative and production-level development environments.
Teaching Format	The course adopts a blended, student-centered approach that emphasizes problem-based learning and active engagement. Selected lecture content is made available online in advance, enabling students to explore key concepts independently and at their own pace. This preparatory work allows in-person sessions to focus on applying knowledge through problem-solving, collaborative activities, and guided discussions—fostering critical thinking and deeper understanding. The teaching format combines frontal lectures, hands-on lab assignments, and project work, ensuring that students develop both theoretical knowledge and practical programming skills. The course is aligned with the principles of the EDUNEXT initiative (https://edunext.eu), promoted by Italian universities, which supports the integration of digital resources and active learning strategies in higher education.
Required Readings	*Python for Data Analysis.* By Wes McKinney. O'Reilly, 3nd Edition, 2022, https://www.oreilly.com/library/view/python-for-data/9781098104023/
Supplementary Readings	Jupyter Notebook Documentation. https://jupyter-notebook.readthedocs.io/en/stable/

Course Module

Course Constituent Title	M2 - Data visualization and exploration
Course Code	27500B
Scientific-Disciplinary Sector	INFO-01/A
Language	English
Lecturers	Prof. Antonio Liotta, Antonio.Liotta@unibz.it https://www.unibz.it/en/faculties/engineering/academic-staff/person/41903
Teaching Assistant	
Semester	First semester

CP	6
Responsible Lecturer	
Teaching Hours	- 24 hours of in-person lectures - 12 hours of video lectures (counted as 24 hours to account for re-watching)
Lab Hours	20
Individual Study Hours	-
Planned Office Hours	18
Contents Summary	This module guides students through the complete data science pipeline, from raw data acquisition to advanced analytics and visualization. Students will gain hands-on experience in data ingestion, exploration, cleaning, and feature engineering, building a strong foundation for effective data modeling. The course covers key machine learning techniques—including clustering, classification, and regression—alongside model tuning, validation, and testing. Emphasis is placed on producing insightful and reproducible visualizations using specialized Python libraries, enabling students to communicate data-driven findings with clarity and impact.
Course Topics	This course equips students with professional skills and knowledge essential for working with large-scale datasets. Through a hands-on and rigorous approach, students will learn to collect, explore, transform, curate, analyze, and visualize data—selecting the most appropriate techniques based on the nature and structure of the data. The course emphasizes the development of insights through a complete data science pipeline: starting from raw data, progressing through data preprocessing and machine learning modeling, and culminating in advanced visualizations. Students will gain practical experience using specialized Python libraries and will learn how to adopt best practices for reproducible, data-driven analysis and research. Common pitfalls in data analysis will be addressed, helping students to critically assess results and avoid misleading interpretations. By the end of the course, students will be able to design and implement robust analytical workflows that support decision-making and scientific inquiry.
Teaching Format	The course adopts a blended, student-centered approach that emphasizes problem-based learning and active engagement.

	Selected lecture content is made available online in advance, enabling students to explore key concepts independently and at their own pace. This preparatory work allows in-person sessions to focus on applying knowledge through problem-solving, collaborative activities, and guided discussions—fostering critical thinking and deeper understanding. The teaching format combines frontal lectures, hands-on lab assignments, and project work, ensuring that students develop both theoretical knowledge and practical data analysis, modelling, and visualization skills. The course is aligned with the principles of the EDUNEXT initiative (https://edunext.eu), promoted by Italian universities, which supports the integration of digital resources and active learning strategies in higher education.
Required Readings	EN - Required readings *Data Visualization. A practical introduction.* Haley. Available online *A layered grammar of graphics*. Wickham. Available online
Supplementary Readings	*Fundamentals of Data Visualization*. Wilke. https://www.oreilly.com/library/view/fundamentals-of-data/9781492031079/ *Visualization Analysis and Design*. Munzer. https://www.oreilly.com/library/view/visualization-analysis-and/9781466508910/ *Data Visualization: Charts, Maps, and Interactive Graphics.* Grant. Amazon *Doing Data Science*. Cathy O'Neil, Rachel Schutt. O'Reilly, 2013, https://www.oreilly.com/library/view/doing-data-science/9781449363871/