# Syllabus

## *Descrizione corso*

| | |
|---|---|
| **Titolo insegnamento** | Programming and Visualisation for Data Science |
| **Codice insegnamento** | 73081 |
| **Titolo aggiuntivo** | |
| **Settore Scientifico-Disciplinare** | |
| **Lingua** | Inglese |
| **Corso di Studio** | Corso di laurea magistrale in Informatica per la Data Science |
| **Altri Corsi di Studio (mutuati)** | |
| **Docenti** | prof. Antonio Liotta,<br>Antonio.Liotta@unibz.it<br>https://www.unibz.it/en/faculties/engineering/academic-staff/person/41903<br>dr. Ozan Kahramanogullari,<br>Ozan.Kahramanogullari@unibz.it<br>https://www.unibz.it/en/faculties/engineering/academic-staff/person/47253 |
| **Assistente** | |
| **Semestre** | Primo semestre |
| **Anno/i di corso** | 1 |
| **CFU** | 12 |
| **Ore didattica frontale** | 80 |
| **Ore di laboratorio** | 40 |
| **Ore di studio individuale** | 180 |
| **Ore di ricevimento previste** | |
| **Sintesi contenuti** | • Introduction to Python and programming environment<br>• Python data structures and programming primitives<br>• Object-oriented programming in Python<br>• Writing structured/reusable code in Python: functions, classes, libraries<br>• Code documentation, testing, version-control, and distribution |

| | |
|---|---|
| | • Advanced Python programming<br>• The data science pipeline: from raw data to advanced analytics and visualization<br>• Data ingestion, exploration, cleaning, and pre-processing<br>• Feature analysis and engineering<br>• Machine learning for data modelling: clustering, classification and regression<br>• Model tuning, validation, and testing<br>• Advanced data visualization |
| **Argomenti dell'insegnamento** | Module 1: Programming in Python<br>The course 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.<br><br>Module 2: Data Analysis and Visualization<br>The course 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. |
| **Parole chiave** | Python Programming; data science; data curation; data modeling; machine learning; data visualization. |
| **Prerequisiti** | Basic programming concepts. Linear algebra. Basic Statistics. |
| **Insegnamenti propedeutici** | |
| **Modalità di insegnamento** | 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.<br>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.<br>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. |
| **Obbligo di frequenza** | Not compulsory. Non attending students have to agree with the lecturer on the modalities of independent study at the beginning of the course. |
| **Obiettivi formativi specifici e risultati di apprendimento attesi** | The course belongs to the type "caratterizzanti – discipline informatiche".<br><br>The course is designed to provide specific professional skills for advanced programming in Python and professional skills and knowledge useful when dealing with large-scale datasets.<br><br>Knowledge and understanding:<br>• D1.1 - Knowledge of the key concepts and technologies of data science disciplines<br>• D1.2 - Understanding of the skills, tools and techniques required for an effective use of data science<br>• D1.3 - Knowledge of principles, methods and techniques for processing data in order to make them usable for practical purposes, and understanding of the challenges in this field<br>• D1.9 - Knowledge of the challenges in the field of man-machine interface and of the methods and techniques for overcoming these challenges<br>• D1.12 - Basic knowledge of the main ethical and social implications of data science<br><br>Applying knowledge and understanding:<br>• D2.1 - Practical application and evaluation of tools and techniques in the field of data science<br>• D2.4 - Ability to develop programmes and use tools for the |

| | |
|---|---|
| | analysis and management of data and related infrastructures<br>• D2.8 - Practical application and evaluation of tools and techniques for data analysis<br>• D2.9 - Design, application and evaluation of technologies and tools for human-machine interaction, data exploration and data visualization<br>Making judgments<br>• D3.2 - Ability to autonomously select the documentation (in the form of books, web, magazines, etc.) needed to keep up to date in a given sector.<br>Communication skills<br>• D4.1 - Ability to use English at an advanced level with particular reference to disciplinary terminology<br>• D4.2 - Ability to present one's work in a clear and comprehensible way in front of an audience, including non-specialists<br>• D4.3 - Ability to structure and draft scientific and technical documentation<br>Learning skills<br>• D5.3 – Ability to deal with problems in a systematic and creative way and to appropriate problem solving techniques. |
| **Obiettivi formativi specifici e risultati di apprendimento attesi (ulteriori info.)** | |
| **Modalità di esame** | The exam modalities are the same for both the attending and the non-attending students.<br>Project work (70% of the final grade) and oral exam (30% of the final grade).<br>All project works must have been submitted, at the very latest, 15 days ahead of the oral exam.<br>In case of a positive mark, the projects will count for all 3 regular exam sessions. |
| **Criteri di valutazione** | 70% project work, 30% oral exam.<br>• 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<br>• Ability to use Python to write, evaluate and deploy advanced, object-oriented computer programs<br>• 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. |
| **Bibliografia obbligatoria** | *Data Visualization. A practical introduction.* Haley. Available online<br><br>*A layered grammar of graphics*. Wickham. Available online<br><br>*Python Data Science Handbook*, by Jake VanderPlas. O'Reilly Media (1st Edition, 2016).<br><br>Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it |
| **Bibliografia facoltativa** | *Fundamentals of Data Visualization*. Wilke. Available online<br><br>*Visualization Analysis and Design*. Munzer. Amazon<br><br>*Data Visualization: Charts, Maps, and Interactive Graphics.* Grant. Amazon<br><br>*Doing Data Science*. Cathy O'Neil, Rachel Schutt. O'Reilly, 2013, https://www.oreilly.com/library/view/doing-data-science/9781449363871/<br><br>*Python for Data Analysis.* By Wes McKinney. O'Reilly, 2nd Edition, 2017, https://www.oreilly.com/library/view/python-for-data/9781491957653/ |
| **Altre informazioni** | Software used: Jupyter Notebook (for Python programming) |
| **Obiettivi di Sviluppo Sostenibile (SDGs)** | Innovazione e infrastrutture, Istruzione di qualità |

# *Modulo del corso*

| | |
|---|---|
| **Titolo della parte costituente del corso** | Programming in Python |
| **Codice insegnamento** | 73081A |

| | |
|---|---|
| **Settore Scientifico-Disciplinare** | INFO-01/A |
| **Lingua** | Inglese |
| **Docenti** | |
| **Assistente** | |
| **Semestre** | Primo semestre |
| **CFU** | 6 |
| **Docente responsabile** | |
| **Ore didattica frontale** | 40 |
| **Ore di laboratorio** | 20 |
| **Ore di studio individuale** | 90 |
| **Ore di ricevimento previste** | |
| **Sintesi contenuti** | • Introduction to Python and programming environment<br>• Python data structures and programming primitives<br>• Object-oriented programming in Python<br>• Writing structured/reusable code in Python: functions, classes, libraries<br>• Code documentation, testing, version-control, and distribution<br>• Advanced Python programming |
| **Argomenti dell'insegnamento** | 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.<br>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. |
| **Modalità di insegnamento** | 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.<br>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. |
| **Bibliografia obbligatoria** | *Python for Data Analysis.* By Wes McKinney. O'Reilly, 3nd Edition, 2022, https://www.oreilly.com/library/view/python-for-data/9781098104023/ |
| **Bibliografia facoltativa** | Jupyter Notebook Documentation. https://jupyter-notebook.readthedocs.io/en/stable/ |

# *Modulo del corso*

| | |
|---|---|
| **Titolo della parte costituente del corso** | Data Analysis and Visualization |
| **Codice insegnamento** | 73081B |
| **Settore Scientifico-Disciplinare** | IINF-05/A |
| **Lingua** | Inglese |
| **Docenti** | prof. Antonio Liotta,<br>Antonio.Liotta@unibz.it<br>https://www.unibz.it/en/faculties/engineering/academic-staff/person/41903 |
| **Assistente** | |
| **Semestre** | Primo semestre |
| **CFU** | 6 |
| **Docente responsabile** | |
| **Ore didattica frontale** | 40 |
| **Ore di laboratorio** | 20 |
| **Ore di studio individuale** | 90 |
| **Ore di ricevimento previste** | |

| | |
|---|---|
| **Sintesi contenuti** | • The data science pipeline: from raw data to advanced analytics and visualization<br>• Data ingestion, exploration, cleaning, and pre-processing<br>• Feature analysis and engineering<br>• Machine learning for data modelling: clustering, classification and regression<br>• Model tuning, validation, and testing<br>• Advanced data visualization |
| **Argomenti dell'insegnamento** | 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.<br>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. |
| **Modalità di insegnamento** | 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.<br>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. |
| **Bibliografia obbligatoria** | *Data Visualization. A practical introduction.* Haley. Available online<br><br>*A layered grammar of graphics*. Wickham. Available online |
| **Bibliografia facoltativa** | *Fundamentals of Data Visualization*. Wilke. https://www.oreilly.com/library/view/fundamentals-of-data/9781492031079/<br><br>*Visualization Analysis and Design*. Munzer. https://www.oreilly.com/library/view/visualization-analysis-and/9781466508910/<br><br>*Data Visualization: Charts, Maps, and Interactive Graphics.* Grant. Amazon<br><br>*Doing Data Science*. Cathy O'Neil, Rachel Schutt. O'Reilly, 2013, https://www.oreilly.com/library/view/doing-data-science/9781449363871/ |