# COURSE DESCRIPTION – ACADEMIC YEAR 2020/2021

| | |
|---|---|
| **Course title** | **Programming and Visualisation for Data Analytics** |
| **Course code** | 73047 |
| **Scientific sector** | ING-INF/05 |
| **Degree** | Master in Computational Data Science (LM-18) |
| **Semester** | 1 |
| **Year** | 1 |
| **Credits** | 12 |
| **Modular** | Yes |

| | |
|---|---|
| **Total lecturing hours** | 80 |
| **Total lab hours** | 40 |
| **Attendance** | Not compulsory. Non attending students have to agree with the lecturer on the modalities of independent study at the beginning of the course. |
| **Prerequisites** | Basic programming concepts |
| **Course page** | https://ole.unibz.it/ |

| | |
|---|---|
| **Specific educational objectives** | The course belongs to the type "caratterizzanti – discipline informatiche" in the curriculum "Data Analytics".<br><br>**Module 1: Data Exploration and Visualization**<br><br>The course is designed to acquire professional skills and knowledge useful when exploring datasets. In particular, the student will be able to visualize datasets choosing the most appropriate technique for the data at hand, and will be able to get insights from the data supported by the visualizations, using basic statistical tools. The student will also learn to avoid the common pitfalls in visualization that can mislead the analysis. Visualization and data handling are done using the R programming language, following the best practices of reproducible research.<br><br>**Module 2: Programming for Data Analytics**<br><br>The course is designed to provide specific professional skills. The students will learn how to organize and analyze data by writing programs. More specifically, the students will practically learn to import, manipulate, analyze, visualize, and model a dataset. The students will also get familiar with libraries that can be effectively used for data analytics. |

| | |
|---|---|
| **Module 1** | **Data Visualisation and Exploration** |
| **Module code** | 73047A |
| **Module scientific sector** | ING-INF/05 |
| **Lecturer** | Matteo Ceccarello |
| **Contact** | mceccarello@unibz.it |
| **Scientific sector of lecturer** | INF/01 |
| **Teaching language** | English |
| **Office hours** | Arranged beforehand by email |
| **Lecturing assistant (if any)** | -- |

| | |
|---|---|
| **Contact LA** | -- |
| **Office hours LA** | -- |
| **Credits** | 6 |
| **Lecturing hours** | 40 |
| **Lab hours** | 20 |
| **List of topics** | • Human psychology and perception<br>• Data and image models<br>• Visualisation software and tools<br>• Visual Diagnostics<br>• Exploratory data analytics<br>• Discovery methods |
| **Teaching format** | Frontal lectures, lab assignments, project work. |

| | |
|---|---|
| **Module 2** | **Programming for Data Analytics** |
| **Module code** | 73047B |
| **Module scientific sector** | INF/01 |
| **Lecturer** | Antonio Liotta |
| **Contact** | antonio.liotta@unibz.it |
| **Scientific sector of lecturer** | ING-INF/05 |
| **Teaching language** | English |
| **Office hours** | Wednesdays 10:30-12:30, to be arranged beforehand by email. |
| **Lecturing assistant (if any)** | -- |
| **Contact LA** | -- |
| **Office hours LA** | -- |
| **Credits** | 6 |
| **Lecturing hours** | 40 |
| **Lab hours** | 20 |
| **List of topics** | • Languages for programming data and data visualisation (Perl, Python, R, Java, Java script)<br>• Programming for text processing (matching, parsing and regular expressions)<br>• Interfacing and integrating (API programming, Plug-in development)<br>• Scripting for data science (e.g., simple shell programming)<br>• Dataflow programming<br>• Advanced Programming Paradigms |
| **Teaching format** | Frontal lectures, lab assignments, project work. |

| | |
|---|---|
| **Learning outcomes** | 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>Applying knowledge and understanding:<br>• D2.1 - Practical application and evaluation of tools and techniques in the field of data science |

| | |
|---|---|
| | • 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. |

| | |
|---|---|
| **Assessment** | Project work (70% of the final grade) and oral exam (30%) |
| **Assessment language** | English |
| **Assessment Typology** | Collegial |
| **Evaluation criteria and criteria for awarding marks** | 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 R software to perform basic data preparation tasks, ability to properly use R plotting facilities, ability to summarize the concepts of the Grammar of Graphics and of human perception, ability to choose the best type of graphical representation for different types of data, correct usage of basic statistical tools<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 and visualization. |

| | |
|---|---|
| **Required readings** | Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it<br>• *Data Visualization. A practical introduction.* Haley. Available online<br>• *R for Data Science*. Wickham. Available online<br>• *A layered grammar of graphics*. Wickham. Available online |

| | |
|---|---|
| | • *Python for Data Analysis.* By Wes McKinney. O'Reilly, 2nd Edition, 2017 https://www.oreilly.com/library/view/python-for-data/9781491957653/ |
| **Supplementary readings** | • *Fundamentals of Data Visualization*. Wilke. Available online<br>• *Visualization Analysis and Design*. Munzer. Amazon<br>• *Data Visualization: Charts, Maps, and Interactive Graphics.* Grant. Amazon<br>• *Doing Data Science*. Cathy O'Neil, Rachel Schutt. O'Reilly, 2013 https://www.oreilly.com/library/view/doing-data-science/9781449363871/ |
| **Software used** | Rstudio https://www.rstudio.com/<br>Jupyter Notebook (for Python programing) |