COURSE DESCRIPTION – ACADEMIC YEAR 2019/2020

Course title	**Programming and Visualization for Data Analytics**
Course code	76050
Scientific sector	INF/01
Degree	Master in Software Engineering for Information Systems (LM-18)
Semester	1
Year	1
Credits	12
Modular	Yes
University	UniBZ

Total lecturing hours	80
Total exercise hours	40
Attendance	**Module 1: Data Exploration and Visualization** Not compulsory. Non attending students have to agree with the lecturer on the modalities of independent study at the beginning of the course. **Module 2: Programming for Data Analytics** Attendance is not compulsory, but non-attending students have to contact the lecturer and the lecturing assistant, at the start of the course, and inform them about it. Both non-attending and attending students have to pass 50% of the lab assessment to be admitted to the written exam. The lab assessment consists of assignments and a project work.
Prerequisites	**Module 1: Data Exploration and Visualization** Basic programming concepts **Module 2: Programming for Data Analytics** The students should be familiar with: - computer programming - linear algebra - statistics
Course page	https://ole.unibz.it/

Specific educational objectives	The course belongs to the type caratterizzanti – discipline informatiche and is part of the Specialization Topics. **Module 1: Data Exploration and Visualization** 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.

	Module 2: Programming for Data Analytics 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 Visualization and Exploration**
Module code	76050A
Module scientific sector	INF/01
Lecturer	Matteo Ceccarello
Contact Lecturer	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
Exercise hours	20
List of topics	• Human psychology and perception • Data and image models • Visualization software and tools • Visual Diagnostics • Exploratory data analytics • Discovery methods
Teaching format	Frontal lectures, lab assignments, project work.

Module 2	**Programming for Data Analytics**
Module code	76050B
Module scientific sector	INF/01
Lecturer	Mehdi Elahi
Contact LA	meelahi@unibz.it Office POS 1.13 https://linkedin.com/in/mehdielahi
Scientific sector of lecturer	INF/01
Teaching language	English
Office hours	Office hours are available by appointment.
Lecturing Assistant (if any)	TBD
Contact LA	TBD
Office hours LA	TBD
Credits	6
Lecturing hours	40
Exercise hours	20
List of topics	• Introduction to programming languages for Data Analytics (Python, R) • Data pre-processing (parsing, normalization) • Exploratory data analysis (clustering, visualization)

	• Building and evaluating predictive models (regression, classification) • Designing the data pipeline
Teaching format	• Frontal Lectures • Lab exercises • Project Work

Learning outcomes	**Knowledge and understanding:** D1.3 To know in depth the scientific method of investigation applied to complex systems and innovative technologies that support information technology and its applications; D1.8 To be able to read and understand specialist scientific documentation, such as conference proceedings, articles in scientific journals, technical manuals. **Applying knowledge and understanding:** D2.1 To know how to apply the fundamentals of empirical analysis of ICT data to the construction of mathematical models for the evaluation and prediction of characteristics of applications and software systems; D2.2 To be able to design and perform experimental analyses of information systems in order to acquire measures related to their behaviour and to evaluate experimental hypotheses in different fields of application, such as business, industrial or research; **Making judgments:** D3.1 To be able to autonomously select documentation from a variety of sources, including technical books, digital libraries, technical scientific journals, web portals or open source software and hardware tools; **Communication skills:** D4.2 To be able to present the contents of a scientific/technical report to an audience, including non-specialists, at a fixed time; D4.3 To be able to structure and draft scientific and technical documentation describing project activities; D4.5 To be able to prepare and conduct technical presentations in English; D4.8 To be able to synthesize knowledge gained from reading and studying scientific documentation. **Learning skills:** D5.1 To be able to independently extend the knowledge acquired during the course of study by reading and understanding scientific and technical documentation in English;

	D5.3 In the context of a problem solving activity, to be able to extend knowledge, even if incomplete, taking into account the final objective of the project; D5.4 To be able to formulate and validate theories and define new methods through empirical induction and new generation scientific investigation tools.

Assessment	**Module 1: Data Visualization and Exploration** Project work in groups and final computer-based exam **Module 2: Programming for Data Analytics** The final mark will be awarded based on: • 50% lab assessment (lab assignments and a project work) • 50% written exam assessment The lab assessment evaluates the capability of the students in applying the data analytics techniques in practical settings. The written exam assessment evaluates the ability of students to understand, recall, and use these techniques, reviewed in the course. In order to be admitted to the written exam, the students have to pass 50% of the lab assessment (i.e., lab assignments and a project work). In case the lab assessment is positive but the final written exam is not positive, the lab grade is valid for three regular exam sessions. The project work is mainly for evaluating the programming skills of the students and it would be the participation in a competition. The students are given a data analytics task, and each of them has to individually develop a program for computing the solution. The students who submit a solution receive the mark for the project work, according to their positions in a leaderboard. For every regular exam session, the complete description of the competition, the rules, and the timeline will be provided in the course page. At the end of the competition, the source code and a 1-2 page(s) final report should be delivered by each of the students.
Assessment language	English
Assessment typology	Collegial commission
Evaluation criteria and criteria for awarding marks	**Module 1: Data Visualization and Exploration** 30% group project work, 70% computer-based exam • Relevant for project work: ability to work in team, clarity of presentation, ability to gain useful and novel insights from data, creativity, critical thinking, ability to adhere to reproducible research best practices • Relevant for written assessment: 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 Non attending students take the same exam as all the other students **Module 2: Programming for Data Analytics** The final mark is composed of: • 50% lab assessment (lab assignments and a project work) • 50% written exam assessment In order to be admitted to the written exam, the students have to obtain 50% of the lab assessment. The lab assessment evaluates the ability of students to apply their data analytics knowledge in practical settings. The written assessment evaluates the ability of students to understand, recall, and use their knowledge (principles and methods) of data analytics.

Required readings	Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it **Module 1: Data Visualization and Exploration** The following required readings are all available online for free • *Data Visualization. A practical introduction.* Haley. Available online • *R for Data Science*. Wickham. Available online • *A layered grammar of graphics*. Wickham. Available online Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it **Module 2: Programming for Data Analytics** • Sebastian Raschka, **Python Machine Learning:** Unlock Deeper Insights into Machine Learning With This Vital Guide to Cutting-edge Predictive Analytics, Packt Publishing Ltd, 2015 (Permanent link at library access to Safari Books: Link)
Supplementary readings	**Module 1: Data Visualization and Exploration** • Fundamentals of Data Visualization. Wilke. Available online • Visualization Analysis and Design. Munzer. Amazon • Data Visualization: Charts, Maps, and Interactive Graphics. Grant. Amazon **Module 2: Programming for Data Analytics** • Brett Lantz, **Machine learning with R**. Packt Publishing Ltd, 2015. (Permanent link at library access to Safari Books: Link)
Software used	Rstudio https://www.rstudio.com/ Jupyter Notebook (for Python programing)