COURSE DESCRIPTION – ACADEMIC YEAR 2018/2019

Course title	**Advanced Data Management Technologies**
Course code	73000
Scientific sector	INF/01
Degree	Master in Computational Data Science (LM-18)
Semester	1
Year	1
Credits	6
Modular	No

Total lecturing hours	40
Total lab hours	20
Attendance	Generally, attendance is not compulsory, but non-attending students have to contact the lecturer at the start of the course to agree on the modalities of the independent study.
Prerequisites	
Course page	https://ole.unibz.it/

Specific educational objectives	The course belongs to the type "caratterizzanti – discipline informatiche" in the curricum "Data Analysis". The course aims at teaching both scientific foundations and practical aspects of advanced data management technologies that go beyond traditional (relational) database management systems. The students will learn the basic concepts of such systems and how to use them to solve concrete problems. Moreover, students will be trained to evaluate the advantages and disadvantages of such technologies in different application contexts.

Lecturer	Johann Gamper
Contact	Piazza Domenicani 3, Room 2.15, gamper@inf.unibz.it, 0471-016140
Scientific sector of lecturer	INF/01
Teaching language	English
Office hours	During the lecture time span: Monday, 13:00-14:00, or email arrangement.
Lecturing Assistant (if any)	Anton Dignös
Contact LA	Piazza Domenicani 3, dignoes@inf.unibz.it
Office hours LA	Email arrangement
List of topics	• Data warehousing and business intelligence • Multidimensional modelling and OLAP • NoSQL and map-reduce • Distributed databases and peer-to-peer systems • Distributed access structures • Main memory database systems
Teaching format	Frontal lectures and project work during the exercise hours. In the frontal lectures, the basic concepts are introduced and explained together with some examples. In the labs, the students will do a semester project, where selected techniques have to be applied to solve concrete problems.

Learning outcomes	Knowledge and understanding: • D1.1 - Knowledge of the key concepts and technologies of data science disciplines • D1.4 - Sound basic knowledge of storing, querying and managing large amounts of data and the associated languages, tools and systems Applying knowledge and understanding: • D2.1 - Practical application and evaluation of tools and techniques in the field of data science Making judgments • 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. Communication skills • D4.1 - Ability to use English at an advanced level with particular reference to disciplinary terminology. • D4.3 - Ability to structure and draft scientific and technical documentation Learning skills • D5.1 - Ability to autonomously extend the knowledge acquired during the study course.

Assessment	The assessment of the course consists of two parts: • a single written exam at the end that covers the entire course material (60% of the mark); • a project which is done during the semester and requires students to solve a concrete problem by using methods and technologies taught in the course (40% of the mark). The written exam is a multiple-choice test and verifies knowledge and understanding of the advanced data management methods and techniques learned during the course. The project verifies whether the student is able to apply advanced data management techniques to solve concrete problems. The project is assessed through a final presentation, demo and project report. A positive project mark is a pre-requisite to be admitted to the written exam; there are no other pre-requisites. Both parts (the written exam and the project) must be positive to pass the exam. The exam modalities are the same for attending and non-attending students.
Assessment language	English
Assessment Typology	Monocratic

Evaluation criteria and criteria for awarding marks	The final exam grade is the weighted average of the project mark (40%) and the mark of the written exam (60%). Criteria for the evaluation of the project: correctness of the solution, complexity of the project, technologies used in the solution, quality of the report and the presentation. Criteria for the evaluation of the written exam: correctness.

Required readings	There is no single textbook that covers the entire course. The course material is collected from various textbooks and research papers, including the following ones: • M. Golfarelli, S. Rizzi. Data Warehouse Design: Modern Principles and Methodologies. McGraw-Hill, 2009. • R. Kimball, M. Ross. The Data Warehouse Toolkit, 2nd edition, John Wiley & Sons. • Tom White. Hadoop: The Definitive Guide, 3rd edition, O'Really. Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it
Supplementary readings	Additional sources will be announced during the course.
Software used	Oracle database, Postgres database, Hadoop MapReduce framework