



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Data Management

2324-1-FDS01Q001-FDS01Q001M

Obiettivi

At the end of the module students will be able to define and realize a complete data pipeline from data acquisition to data storage (relational or not) according to their application needs

Students will be able to evaluate for each step the appropriate tool to use

Contenuti sintetici

Introduction to data management in big data context

data life cycle
data acquisition techniques
data modelling and storage
data preparation, cleansing, quality and explorative data analysis
advanced data management concepts

Programma esteso

1. Data life cycle
2. data acquisition
3. web scraping
4. rest api
5. real time data acquisition
6. use of LLM

7. **data storage and modelling**
8. Introduction to NoSQL models
9. key value and columnar models
10. Document based system
11. Graph db
6. **data preparation, cleansing, quality and explorative data analysis**
12. Data integration
13. Data quality
14. **Advanced data management concepts**
15. data warehouse
16. data lake

Prerequisiti

knowledge of relational model

Modalità didattica

Lectures and exercises in the classroom and on virtual lab

Lectures with the support of slideware, discussion of practical cases through the forum, discussion of practical home-work projects.

Some self-assessment tests, not considered for the final evaluation will be provided

Materiale didattico

G. Harrison Next Generation Databases, Apress, 2015

A. Rezzani Big data analytics Apogeo 2017

Yau, N. (2011). *Visualize this: the FlowingData guide to design, visualization, and statistics*. John Wiley & Sons.

Ware, C. (2012). *Information visualization: perception for design*. Elsevier.

Scientific articles and class pack provided by the lecturers.

Periodo di erogazione dell'insegnamento

first semester

Modalità di verifica del profitto e valutazione

The exam is divided into two parts

Data Management (50% of the final evaluation): Written exam and a project related to the topic of the module

Data visualization(50% of the final evaluation): test and a project related to the topic of the module

Orario di ricevimento

Please send an e-mail to teachers to arrange an appointment

Sustainable Development Goals

ISTRUZIONE DI QUALITÀ | IMPRESE, INNOVAZIONE E INFRASTRUTTURE
