UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Introduction To Databases

2627-2-E4103B012

Learning objectives

The course is an introduction to databases, with a focus on relational databases. Database design methodologies will be illustrated. The course helps consolidating the knowledge and understanding of the computer science area in the context of the degree in Statistics and Economics Science. The relational model and the SQL language will be illustrated. It will be taught how queries can be used to compute synthetic indicators starting from information stored in relational databases. The student will be able to understand the structure of an existing relational database, design a new database, and formulate queries to extract data that answers questions. During the laboratory activities, students will be encouraged to interact with one another to improve their teamwork skills. The student will also acquire the knowledge and skills to independently learn other database technologies, e.g., NoSQL databases.

Contents

Course topics

- Introduction to relational databases
- The Client / Server paradigm
- Relational algebra and SQL language
- Analysis and extraction of information from existing databases
- Conceptual, logical, and physical design of databases
- Transaction support
- Structured, semi-structured, unstructured data

The teacher might change the topics taught. Any changes will be promptly indicated on the course website.

Detailed program

- The Relational paradigm
- Primary Key and Foreign Key
- Client / Server paradigm
- Centralized data management
- Relational algebra (projection, selection, cartesian product, join)
- SQL Language (Data Definition and Data Manipulation Language)
- Anatomy of an SQL query
- Group By clause and aggregation operators
- Nested Queries
- Calculation of indicators through SQL query
- Information Systems Lifecycle and Database Design
- Conceptual design
- Logical design
- Short notes on Physical design
- ER Model
- Normalization
- Transactions and ACID properties
- Indexes
- Structured, semi-structured, unstructured data

Prerequisites

None. It is highly recommended to have previously attended an introductory computer science course where the following concepts were introduced.

- Boolean algebra (especially the AND, OR and NOT operators)
- Variables
- Conditional statements
- File System

Teaching methods

This is a blended-elearning course and the teaching activities will be performed in several ways:

- Frontal lessons, live and in person or video-recorded
- E-learning lessons (recorded lectures, quizzes or other activities to elaborate the contents, self-assessment exercises)
- Exercises in the computer lab
- Final exercise with simulation of the learning assessment methods

About 52% of the lessons (approximately 23 hours) will be delivered in the form of "traditional lectures", through frontal and e-learning lessons.
About 48% of the course (approximately 21 hours) will be delivered through "interactive teaching" in the form of laboratory exercises.

Some of the frontal lessons hours might be replaced by testimonials and expert dialogue sessions. In this case, the testimonials will take place approximately in the last weeks of class, based on the availability of the people

involved. Dates and times of these events will be communicated on the course website.

The laboratory activities will take place in person in computer labs. If computer labs are unavailable, these activities may also take place in traditional classrooms (electrified or not) where students can use their personal PCs, or in remote virtual labs.

Assessment methods

The learning assessment does not include ongoing tests but only the exam after the end of the course (available in different sessions, according to the university regulations). The teacher will discuss in class whether to introduce ongoing tests at the beginning of the course, and if so, a communication will be sent to students via the course website.
Students must register before taking the exam. It is mandatory to register using the University Information System by the official deadline. Registration usually closes 3 working days before the examination day.

How Learning Verification is performed

The learning verification of the Databases Course ("exam" from here on) is composed of a written test and a subsequent discussion/acceptance of the grade. The former is mandatory, the latter is optional (at the discretion of the teacher and the student).
The teacher, at his discretion, may require a student to take the oral exam if there are doubts about the written exam.

The examination is the same for both attending and non-attending students.

The written test consists of both closed-ended and open-ended questions; the number and types of questions may change. The written exam duration may change based on the number of exercises to be completed and the test's complexity. As a general reference, you can expect a duration of 1 and a half hours.

For each closed-answer question, you will be offered a set of possible answers, and you will have to choose the correct answer (one and only one will be the correct answer, unless otherwise indicated in the assignment). Wrong answers will not result in penalties; correct answers will positively contribute to your final grade. Closed-ended questions will focus on theory topics or require completing a query or identifying the output produced by a query. Other closed-ended question types may be added. A mock exam will be available at the end of the course, allowing students to take an online exam simulation. The text of the simulation and thesolutions will be available on the course website.

In the open-ended questions, you may be asked to:

- Write a query in SQL language. You will not have a database (management system) for testing your query.
- Design the ER schema of a database.
- ... (other types of exercises focusing on topics presented in class)

At the end of the course, after the exam simulation, a section of the course website containing several previous exam tests and the corresponding solutions will be made available.

To pass the exam, you will have to reach a sufficient level in both the closed-ended questions and the open-ended questions. I.e., you will need to demonstrate adequate theory knowledge and be able to solve exercises (e.g., write an SQL query, design a data structure). During the exam, each student will have a personalized assignment (with questions that are partially or totally different from those of all other students). Regarding the ability to write SQL queries, you can use a simple test to check your preparation: if, at the end of the course, you need to look at the solutions to solve any of the exercises proposed during the lab activities, then you need to study and practice more.

In the event of a seriously insufficient examination, there are no limitations on registering for the next exam sessions. The teacher trusts in students' collaboration; however, he reserves the right to introduce limitations in the future if repeated and widespread inappropriate behavior occurs. It is considered inappropriate behavior a) to show up for an exam session completely unprepared or b) not withdrawing in case of a severely insufficient exam. The introduction of any limitations, if any, will be publicized on the course website.

The written exam is held in Italian. Students who do not speak Italian may agree with the teacher to replace the written exam with an oral exam in English or German.
.

Textbooks and Reading Materials

Study material

- The slides, exercises and all the material showed in class will be published on this website.
- Additional material may be shown during the course. The references will be published on this website.

Book

- Atzeni, P., Ceri, S., Paraboschi, S., Torlone, R. Databases. McGraw-Hill. Any edition of the book is ok.

Semester

First half of the first semester.

Teaching language

Italian

Sustainable Development Goals

INDUSTRY, INNOVATION AND INFRASTRUCTURE