

Degree	Type	Year
2503758 Data Engineering	OB	2

## Contact

Name: Enric Marti Godia

Email: enric.marti@uab.cat

## Teachers

Enric Marti Godia

## Teaching groups languages

You can view this information at the [end](#) of this document.

## Prerequisites

It is recommended that the student have the knowledge and skills of:

- Programming in third generation languages (C, PHP, Java, etc.).
- Basic data structures.

## Objectives and Contextualisation

In this subject, the basic concepts of Relational Databases (BDR) necessary for both BD and user designers are introduced.

KNOWLEDGE: At the end of the course the student must be able to:

- Understand and understand the important technological leap that represent the databases and BDR systems in particular, with regard to computer information processing, as well as in the design and maintenance of information processing applications.
- Know the architecture of the BDR systems, the functions of each module and the personnel working on these systems (users, programmers and administrators of BD). Architecture is studied from a remote and local point of view.
- Study the properties of the relational model of BD, extended in most BD engines.
- Know and know how to apply the SQL language, standard in BDR.
- Know and know how to apply the Model Entity / Relationship (E / R), very used in BDR design.
- Understand the design methodology of BD, including normalization techniques of a BDR.
- Know the main data storage structures that are used in BDR, such as indexing and hashing functions.

ABILITIES: It is intended that students acquire the following abilities:

- Use the Rules of Integrity of the relational model of BD.
- Perform simple queries and a certain complexity in a BDR using SQL.
- Design a BD in the Model E / R based on real-world specifications.
- Convert the BD into the E / R Model to a set of relationships and attributes of a BDR, applying standardization techniques.
- Work with an example of a BDR engine such as ORACLE, which is widely used in the professional field, at the user level.

## Competences

- Conceive, design and implement efficient and secure data storage systems.
- Handle large volumes of heterogeneous data.
- Prevent and solve problems, adapt to unforeseen situations and take decisions.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

## Learning Outcomes

1. Design and configure data-processing infrastructures on the basis of particular requirements.
2. Design relational or non-relational databases suited to the data to be processed.
3. Perform queries on databases.
4. Prevent and solve problems, adapt to unforeseen situations and take decisions.
5. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
6. Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.
7. Write technical reports on a database.

## Content

### BLOCK 1. DATABASE PARADIGMS

#### 1. Introduction. Basic Concepts (2 hours)

- Introduction and definitions
- Components of a Database System
- Historical Evolution
- Advantages and disadvantages of a database system

#### 2. Architecture of Relational Databases (2 hours)

- ANSI-SPARC architecture
- EI DBA and the DBMS
- *Back-end / front-end* architecture

### BLOCK 2. RELACIONAL MODEL

#### 3. RELational Model Relacional: Data Structure and Integrity Rules (2 hours)

- Data Structure
- Integrity Rules

#### 4. Relacional Model: Data Manipulation (6 hours)

- Relacional Algebra
- Càlcul Relacional Calculus

### BLOCK 3. DATABASE DESIGN

#### 5. Conceptual Design (8 hours)

- Design phases of a Database
- Capture and analysis of requirements
- E/R Diagram
- Extended E/R
- Design Criteria of an E/R scheme

#### 6. Logical Design (2 hours)

- Conversion to tables (Relacional Model)

#### 7. Normalization (2 hours)

- Normalization Theory
- Normal Forms: 1NF, 2NF, 3NF
- Boyce-Codd Normal Form (BCNF)

### BLOCK 4. PHYSICAL LEVEL

#### 8. Physical Level (2 hours)

- Access to the physical database. Files
- Storage structures
  - Indexing
  - Hashing
- Compression techniques

### Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Explicación y resolución de problemas	26	1.04	3, 4, 6, 7
Master classes	26	1.04	1, 2, 5
Type: Supervised			
Technical Rapport Feedback	2	0.08	4, 6, 7
Type: Autonomous			
Estudi individual	26	1.04	1, 2, 5
Realització problemes SQL i disseny de BDR	52	2.08	1, 2, 3, 4, 6

LEARNING MANAGEMENT SYSTEM (LMS): Caronte (<http://caronte.uab.cat>, degree in Enginyeria de Dades, subject *Bases de Dades Relacionals*, course *BDR Curs 2024-25*).

The student's learning process will be based on these two types of classroom activities: Theory and Seminars and an autonomous work activity, the Technical Report. Attendance at theory classes, and seminars IS NOT COMPULSORY, BUT YOU ARE VERY RECOMMENDED.

## THEORY

The theory classes are taught through magisterial classes with transparencies and slate. All the material of the course will be accessible through the LMS. Some of the theory sessions will be devoted to solve exercises in the same classroom with the aim of deepening the most relevant concepts and teaching the student to identify and analyze the typical errors.

## SEMINARS

It consists of two parts. In the first part the SQL language will be introduced and to practice an SQL script is provided on a BDR with content and a statement with 120 queries that the student must solve. The resulting tables are provided so that the student can practice outside the classroom. In the seminar sessions, consultations are proposed to be solved, which the students do with the teacher's tutoring. An account is provided to the Oracle server of the Engineering School and the link to download the Oracle sqlDeveloper client software so that students can send their queries to the Oracle server of the School. In a more local way, the link is provided to download a reduced version of the Oracle (Oracle Express Edition) engine, to be able to be installed on a PC or laptop in a local modus.

In the second part of the course, a statement with design problems is disproportionate, some of them, at the teacher's proposal, students solve in class with their support.

All documentation (statements of SQL and design, BD scripts, results of problems) of seminars is available to the LMS.

## TECHNICAL RAPPORT

An autonomous and cooperative work activity is proposed: case of use in the design of a real database.

In this activity a real case of design of BD will be presented to the student so that the student performs all the phases of design: capture and analysis of requirements, design E / R, logical design and game of SQL queries. This activity will be done in groups of 5 people that must be registered via LMS. The different stages of the technical report will be monitored throughout the course. There will be two deliverables evaluable in LMS: a first delivery of the design E / R with which to obtain the teacher's feed-back and final delivery at the end of the semester with all stages.

Cooperative work is carried out both in the design activity of a BD in which the student will have to work in a group of 5 people and in carrying out the problems proposed in the seminar sessions.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

## Assessment

### Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
-------	-----------	-------	------	-------------------

SQL Exam	25%	2	0.08	3
Technical Rapport	25%	2	0.08	4, 6, 7
Theory Exams	50%	4	0.16	1, 2, 5
Voluntary exercise delivery	+1,0	2	0.08	2, 3, 6

The evaluation of the subject will be carried out continuously, based on three indicators: Theory, SQL and the Technical Report.

## THEORY

There will be two individual written exams, with a weight of 50% each on the Theory Note. The first test (Par1) will be done approximately halfway through the semester and will evaluate the theoretical concepts explained up to date and the ability to solve BDR queries in relational algebra. The second test (Par2) will be carried out at the end of the semester and will evaluate the concepts explained until the time of theory (design of BD, normalization, physical level) and the ability to solve assumptions of design E/R and step to relational model.

Recovery test: In case the theory note does not reach the approved, students may submit to a review of recovery on the contents of the partial examination suspended.

Seminar classes offer two voluntary deliveries valued each with 0.5 points to add to the partial exam mark ( Par1 or Par2). The first delivery consists of SQL query exercises where the student's abilities are evaluated in the resolution of SQL BDR queries (NPrb1). The second consists of a design exercise of BD, where based on some requirements, the most appropriate E/R diagram must be designed (NPrb2).

## SQL EXAM

Individual examination by computer where the student's abilities are evaluated in the resolution of SQL queries. A series of random SQL queries are proposed on the BD used in the seminars that the student must solve in a given time (SQL Note).

Recovery Test: If the SQL exam is not reached on 5, the student has a new opportunity with the same format and rules

## TECHNICAL RAPPORT

The Technical Report is an assessable group activity consisting of 4 parts, corresponding to the 4 design phases of a BDR: Requirements, E / R Diagram, Logical Design in tables and Implementation of a set of SQL queries. A written document that contains the work carried out in the 4 phases, duly justified, will be presented. The document must have the appropriate format so that the diagrams of the E / R Diagram and the corresponding Relational Diagram can be read correctly. The Technical Report Note will be obtained from two deliveries: The first one, with the writing of the data requirements of the presented case and the E / R Diagram, resulting from the design of the BD. The second installment will contain the same points of the first one including the teacher's corrections by adding the Relational Diagram (resulting from the logical design) and the code of the SQL queries and their results in tables. The note of the Technical Report will be the weighted average of the note obtained from the two deliveries: Design E / R (60% of the note) and the Logical Design with SQL queries (40%).

If a diagram of the Technical Report could not be read clearly for reasons of resolution or quality, this part of the report will be suspended, as well as the parts that depend on it as it can not be evaluated.

## QUALICATION INDICATORS

The final grade of the subject is calculated in the following way (between square brackets the minimum marks to make half):

- $Partial\ Mark\ \#1 = (Par1 + NPrb1) [Par1 \geq 4,5]$
- $Partial\ Mark\ \#2 = (Par2 + NPrb2) [Par2 \geq 4,5]$
- $Theory\ Mark = 0'5 * (Partial\ Mark\ \#1) + 0'5 * (Nota\ Partial\ Mark\ \#2) [Partial\ Mark\ \#i \geq 5, i=1,2]$
- $SQL\ Mark [SQL\ Mark \geq 5]$
- $Technical\ Rapport\ Mark = 0'6 * E/R\ Design\ Mark + 0'4 * Logical\ Design\ Mark [E/R\ Design\ Mark, Logical\ Design\ Mark \geq 5]$
- $FINAL\ GRADE = 0'5 * Theory\ Mark + 0'25 * SQL\ Mark + 0'25 * Technical\ Rapport\ Mark$

## EVALUATION CRITERIA

- To count the marks of the voluntary deliveries of the seminars (NPrb1 and NPrb2) you must obtain a minimum of 4.5 in the corresponding partial exam note (Par1 or Par2).
- In order to calculate the FINAL SUBJECT, we will have to pass all the evaluable activities (examinations of Theory, SQL Exam and the Technical Report) with a 5.
- In all recovery activities (partial exams or internships), the maximum mark will be 8.
- NOT ASSESSABLE: A student is considered Non-Appraiser (NA) only if he has not done any evaluation activity. We recall that the non-valuable note also runs a call.
- NOT PASS: In the case of not reaching the minimum required in any of the assessment activities, if the calculation of the final grade is equal to or greater than 5, a 4 note will be placed on the file.
- MATRICULA DE HONOR (MH) - HIGH DISTINCTION: Students with a mark equal to or greater than 9 points can obtain a MH, at the discretion of the teacher. As the number of MH can not exceed 5% of students enrolled, students who have the highest final grades will be awarded, at the discretion of the teacher.
- REPEATERS: No approved part is approved (Theory, SQL Exam, Technical Report) from one academic year to another.
- IMPORTANT FOR ALL STUDENTS: It is important to enroll in Caronte (<http://caronte.uab.cat>, subject *Bases de Dades Relacionals*, curs *Docència BDR curs 2019-20*) at the beginning of the course, because the materials are published of the subject, the deliveries of the Technical report are made and the final notes of the subject are published. If this is the first time you enter Caronte, you must provide your NIU, password, first and last name, email and a photo ID in JPG format. Caronte is a manager different from the UAB Virtual Campus, which means that passwords mustn't be the same.

## EVALUATION CALENDAR:

- Partial Exams: Dates at the end of the teaching guide published in the LMS and provided to students on the first day of class.
  - Recovery Exams: According to the academic calendar of the School of Engineering.
  - Voluntary deliveries in the seminar sessions: Noticed in advance, since being a voluntary activity, it is intended to encourage attendance at seminars throughout the semester.
  - Delivery Technical Reports: Dates at the end of the teaching guide provided to the students on the first day of class, appearing in the course of Caronte the corresponding week.
- CALENDAR OF EVALUATION:

The dates of continuous evaluation will be published in Caronte and may be subject to changes of programming for reasons of adaptation to possible incidents. Charon will always be informed about these changes as it is understood to be the usual platform for the exchange of information between teachers and students outside the classroom.

Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the current academic regulations, irregularities committed by a student that may lead to a variation of the qualification will be classified by zero (0). Assessment activities qualified in this way and by this procedure will not be recoverable. If it is necessary to pass any of these assessment activities to pass the subject, the aquatic will be suspended directly, without opportunity to recover it in the same course. These irregularities include, among others:

- The total or partial copy of a practice, report, or any other evaluation activity.
- Let cop.
- Present a group work not done entirely by the members of the group.

- Present as own materials produced by a third party, although they are translations or adaptations, and generally works with non-original and exclusive elements of the student.
- Have communication devices (such as mobile phones, smart watches, etc.) accessible during theoretical-practice tests (individual exams).

With the exceptions that the "non-evaluable" qualification will be awarded to students who do not participate in any of the assessment activities, and that the numerical note of the file will be the lowest value between 3.0 and the average Weighted notes in case the student has committed irregularities in an evaluation act (and therefore not approved for compensation).

SUMMING UP:: copying, copying, or plagiarizing in any of the assessment activities is equivalent to a SUSPENS with a score of less than 3.5.

## Bibliography

SUBJECT MATERIALS (docs, slides, papers): In the LMS.

BASIC BIBLIOGRAPHY:

- A. Silberschatz, H.F. Korth, S. Sudarshan, Fundamentos de Bases de Datos, 5a edición, McGraw-Hill, 2006.
- R. Elmasri, S.B. Navathe, Fundamentos de Bases de Datos, 5ª edición, Addison Wesley, 2007.

COMPLEMENTARY BIBLIOGRAPHY:

- T.M. Connolly, C.E. Begg, Sistemas de Bases de Datos, 4a edición, Addison Wesley, 2005.
- C.J. Date, Introducción a los sistemas de Bases de Datos, 7a edición, Prentice Hall, 2001.
- A. Opper, R. Sheldon, SQL, McGraw-Hill, 2009.
- A. Armendáriz, T.A. Pérez, C. Presedo, M.Azanza, Casos prácticos para diseño de Bases de Datos, 2019.
- S. Garrido, Diseño de Bases de Datos. Un enfoque práctico, 2019.
- M. Celma, J.C. Casamayor, L. Mota, Bases de Datos Relacionales, Pearson-Prentice Hall, 2003.
- D.M. Kroenke, Procesamiento de Bases de Datos, 8ª edición, Pearson-Prentice Hall, 2003.
- A. de Miguel, M. Piattini, Diseño y uso de Bases de Datos Relacionales, Ra-Ma, 1997.

WEB LINKS:

- <https://oai.oracle.com/> , Oracle Academic Initiative (OAI) with a lot of information concerning to courses and activities about Oracle for UAB students (last access: july 2024).
- <http://ilearning.oracle.com/ilearn/en/learner/jsp/login.jsp>, Oracle online free and to pay courses. (last access: july 2024).
- <http://www.acm.org/sigmod> Special Interest Group in Management of Data. Grup de l'ACM (Association of Computer Machinery) that promote a lot of activities about DataBases, they organize international conferences and journals about DataBases (last access: july 2024).

MULTIUSER RELATIONAL DATABASES:

- <http://www.oracle.com/> , Oracle® (last access: july 2024).
- <http://www.mysql.com/>, MySQL® (last access: july2024).
- <http://www.postgresql.org>, PostgreSQL® (last access: july 2024).
- <http://www.sybase.com/home> , Sybase® (last access: july2024).

- <http://www.microsoft.com/sql/default.asp>, Microsoft SQL Server® (last access: july 2024).
- <http://www-4.ibm.com/software/data/db2/>, IBM DB2® (last access: july 2024).
- <http://www-01.ibm.com/software/data/informix/>, IBM Informix® (last access: july 2024).

## Software

In the seminar part we will use the following software:

- Oracle. Database engine and server. There is a server at the School of Engineering, from which students will receive an access code and instructions on how to access it.
- You can also download a reduced version of the Oracle server (Oracle 18c Express Edition) which you can download in Windows or Linux at the link: <https://www.oracle.com/database/technologies/xe-downloads.html> . You will be asked to register on the Oracle website.
- Sqldeveloper v.23.1.1 client software to be able to log in to Oracle servers. You can get the software in Windows, MacOS and Linux versions at the link: <https://www.oracle.com/tools/downloads/sqldev-downloads.html> . You will be asked to register on the Oracle website.

All the software that will be used is free access or they are servers of the School to which the students enrolled in the subject will be given access.

## Language list

Name	Group	Language	Semester	Turn
(PAUL) Classroom practices	81	Catalan	first semester	morning-mixed
(PAUL) Classroom practices	82	Catalan	first semester	morning-mixed