

DATABASE PROJECT 23/24

This project will provide students with hands-on experience designing and implementing an SQLite relational database. To do so, students will need to create the conceptual model for the selected theme, map this model to a relational schema, implement the schema in an SQLite database, and populate the database. Additionally, it incorporates the innovative use of generative AI tools to assist in various stages of the database development process.

1 IMPORTANT DATES

- Working Group Definition [September 24, 11:55 PM]
- Problem Domain Definition [October 1, 11:55 PM]
- First Submission [October 15, 11:55 PM]
- Second Submission [November 12, 11:55 PM]

2 WORKING GROUP DEFINITION

Students will do the project in groups of 3. Groups should be defined in Moodle's "Groups for the project" activity. The first digit of the group identifier should be the class number. For example, Group 101 should be composed of students from class 2LEIC01.

3 PROBLEM DOMAIN DEFINITION

Students shall propose a domain/topic for the database implementation to the lecturer of the theoretical-practical classes. After the approval of the topic, the group must send, through Moodle, a description of its subject with a maximum of 100 words. This description should state what type of application could use the database and the information it will contain.

For your reference, we expect relational schemas with 10 to 15 relations. Some of the relations should have composite keys.

4 FIRST SUBMISSION: CONCEPTUAL MODELLING

In this submission, students will focus on the initial stages of database development, including conceptual model design and generative AI integration for this task.

4.1 TASKS

A. Domain Definition

1. Familiarise yourself with the context associated with the topic of the work. Understand in detail the data that the database must store.
2. Describe the context, including all information that may be important for evaluating the conceptual model. This description ***should not*** be a description of the conceptual model.

B. Conceptual Modelling

1. Define a conceptual model in UML for a database for the defined problem domain without using generative AI integration.
2. Present an initial solution for the conceptual model. Remember the constraints and associations' multiplicities. If used, derived elements should be indicated.
3. Propose a final solution for the conceptual model based on AI tool integration.
4. Discuss the refinements made after AI tool integration.

C. Generative AI Integration

1. For the conceptual modelling task, use a generative AI tool to assist in solving the problem.
2. Provide a clear description of the specific generative AI tool used and the sequence of prompts given to the tool.
3. Critically assess the output generated by the AI tool, highlighting strengths and limitations.

4.2 DELIVERABLES

- A report in pdf format with one section for each task: A2, B2-4, C2-3.

5 SECOND SUBMISSION: RELATIONAL MODELLING, DATABASE CREATION AND DATA LOADING

5.1 TASKS

A. Refine the conceptual model

Refine the conceptual model based on the feedback you received for the first submission. The refined model should be included in the report even if no changes are needed.

B. Define the relational schema

1. Transform the conceptual model into a relational schema, which should be added to the report in a textual format using the syntax: "R1 (atr1, atr2, atr3->R2)". A clear indication of each relation's primary and foreign keys is expected.
2. Present an initial relational model for the given conceptual model.
3. Propose a final solution for the relational model based on AI tool integration.
4. Discuss the refinements made after AI tool integration.

C. Functional Dependencies and Normal Forms Analysis

1. For each relation, identify its functional dependencies.
2. For each relation, identify the existing violations of the Boyce-Codd Normal Form. Students must justify the non-existence of infringements.
3. Identify the existing violations of the 3rd Normal Form for each relation. Students must justify the non-existence of infringements.
4. Relations that are neither in the Boyce-Codd Normal Form nor the 3rd Normal Form must be decomposed into one of these normal forms.
5. Propose a final solution for the above tasks based on AI tool integration.
6. Discuss the refinements made after AI tool integration.

D. SQLite Database Creation

1. Create a file named create1.sql that includes the SQL statements to create all the relations of the final relational schema designed in Task 2. Before the create table statements, make sure to drop existing relations with the same name. SQLite allows you to read commands from a file. This feature should be used to (re)create the database whenever necessary.
2. For each constraint identified in the refined conceptual model, decide how it should be implemented in SQL and update the create1.sql file to include the SQL statements that implement the constraints. Note that some constraints might require triggers and, in these cases, should not be implemented. It is also necessary to consider that implementing constraints in SQLite is not fully compliant with the SQL-99 (SQL2) standard.
3. Propose a final solution for the create2.sql script based on AI tool integration.
4. Discuss the refinements made after AI tool integration.

E. Data Loading

1. Create a file named populate1.sql that includes the SQL statements to introduce data in all previously created tables. At the beginning of this file, you must have the statement "PRAGMA foreign_keys = ON;" (without quotes) to ensure that the referential integrity check of the database is active.
2. Propose a final solution for the populate2.sql script based on AI tool integration.
3. Discuss the refinements made after AI tool integration.

F. Generative AI Integration

1. For Tasks B-E, use a generative AI tool to assist in solving the problem.
2. Provide a clear description of the specific generative AI tool used and the sequence of prompts given to the tool.
3. Critically assess the output generated by the AI tool, highlighting strengths and limitations.

5.2 DELIVERABLES

- A report in pdf format with one section for each task: A, B2-4, C1-6, D4, E3, F2-3.
- create1.sql, create2.sql, populate1.sql and populate2.sql files.

6 EVALUATION

The overall grade of the project will be computed as:

$0,4 * 1^{\text{st}} \text{ submission grade} + 0,6 * 2^{\text{nd}} \text{ submission grade}.$

6.1 PARTICIPATION OF THE VARIOUS STUDENTS OF THE GROUP

The report for each submission should include a paragraph that qualitatively assesses each group member's contribution. This paragraph will allow us to detect abnormal cases and act accordingly.

6.2 DELAYS

Late submissions will be penalised by one value for each day of delay. No submissions will be accepted later than 1 week from the submission date.

6.3 AUTHORSHIP AND ORIGINALITY OF THE WORK

All works will have their originality thoroughly scrutinised. The authors of prevarications will be punished in cases of similarities with third-party works (works not cited, other works by students of the curricular unit, etc.), which can range from the cancellation of the enrolment in the curricular unit to the initiation of a disciplinary process for all students of the group.