Data Mining and Knowledge Extraction 2023-24
Assignment 1

Preliminaries

Read everything carefully. Make sure that the names of the files are exactly as described here. The goal of this phase is to verify your ability in designing a relational database from the conceptual design (E/R diagram) to the logical design (relational model).

The assignment is done individually (1 person delivers one assignment).

Deadline

This phase has a strict deadline on **Sunday 19th November 2023 at 23:59**. Any assignment delivered after that time will not be considered. Any clarification questions need to be asked on KIRO **before the 17th of November.** No questions will be answered after that date.

Task Description & Marking

You have received a letter from the director of the National Library. You can see the letter in the next page. He is asking you to help design a database they need at the national library, by providing them the E/R diagram of the database, and the SQL DDL commands to create the schema of the relational database.

The Database description is divided into **individual items**. Each item gets:

- 1 point if correctly modeled in the conceptual design (E/R) and another point if correctly modeled on the logical design (SQL DDL), so in total each item can get 2 points.
- If an item is modeled correctly in the E/R but not in the logical design, you get 1 point only.
- If an item is not modeled correctly in the E/R, but your logical design is correct with respect to what you specified in the E/R for that item, you get 1 point only.
- If an item is not modeled correctly in the E/R and in the logical design, you get 0 points for that item.

Delivery

1. Create your ER diagram using some computer tool (draw.io, PowerPoint, dia, yEd, lucichart, …) or anything else you like. No handwritten assignments are accepted.
2. Save your ER diagram, together with the SQL DDL commands in a **single** pdf file. Make sure that your ER diagram fits in **one** page.
3. **The name of the pdf file should be "A1_" followed by your matricola.** For example, if your matricola is 12345 then the file should have the name `A1_12345.pdf`.
4. To deliver, please go to page https://forms.gle/an3ttqxvkC4p6Btc6, **you must login with your UNIPV credentials.**
5. Fill up the fields and upload your pdf file.
6. In case you find a mistake in your assignment before the deadline, you can resubmit an updated version **UP TO A MAXIMUM OF 10 files.**

Giuseppe Librano
Director of the National Library
Lungotevere Ripa, 1
00153 - Roma

Dear student of the Data Mining and Knowledge Extraction course

As the director of the National Library, I must guarantee that all the information regarding our Library is kept well-organized, considering the large amount of material and users that interact with us. For this reason, we would like to restructure all our internal databases to keep track of different information at all levels that might be relevant to our Library. Thus, we would like to ask you as an expert in databases to provide us with an E/R diagram that models the following requirements, together with the SQL DDL commands to create the relational schema. So here is what we have:

1. The database should contain information about books. Each book is characterized by the ISBN. We also need to store the title, year of publication, the publisher and the authors of the book
2. Each book is written by one or more authors.
3. For each author, we must record their names, birthday date, and nationality.
4. A book can also be part of a collection, of which we need to store the name and date of publication.
5. A collection is published by a publisher, and we can identify a collection only when both its name and the publisher that publishes it are known (there can be collections with the same name, but published by different publishers). In particular, we are not interested in keeping information about collections, when the publisher is deleted from our systems.
6. Regarding publishers, we identify them with their Partita Iva (i.e., the fiscal code), and we need to store their name and Head Quarter address.
7. We also want to keep track of all the customers of our Library. In particular, each customer has a registration number, and we want to store their name, surname and birthday date as well.
8. We allow our customers to rent books for a certain amount of time, delimited by a starting date and an ending date. They can of course rent different books. If a customer rents the same book multiple times, for different dates, we are only interested in tracking the last time.
9. Regarding our employees, *some* of them can have a special role: "front office" (i.e., dealing with customers) and "back office" (i.e., ordering books, etc.).
10. All our employees have a social security number that identifies them, and we also want to store their names, surnames, and date of birth.
11. Employees for the front office also have a desk assigned. Moreover, front officers are the only ones that can authorize the renting of books for a customer (every rent must be authorized by precisely one front officer). So, we also need to keep track of the specific front officer that authorized a rent, and the date of the authorization.
12. Regarding employees at the back office, they are mostly in charge of ordering books from our dealers. Dealers are identified by an id, and we want to store their name and address.
13. We must store information about orders that a back officer places for a certain book, at a certain dealer. For such an order, we want to store the order date, and the amount of copies requested for that book. If the same back officer orders the same book from the same dealer on different dates, we want to *keep the history* of all such orders in our system. Assume an order is identified by a unique id.

Sincerely
Giuseppe Librano