Database Systems Course (2021-2022) – HW3

The goal of the project: design and implement a useful and interesting web application related to movies. The application should provide functionality based on your desired target audience – for example, movie producers might be interested in analyzing the characterizations of blockbuster movies (e.g. popular actors and their combination, directors, script writers, length, music), while fans of specific genre would like to know in which movie their favorite actors participated. The possibilities are endless – be creative!

Project objectives:

- Design and create a demo movie-related web-application, with an underlying MySQL database schema, where the information originates from **at least two sources:**
    - One of the sources should be accessed via an API (there are many movie-related APIs online – contact me if you're having trouble finding one)
    - The other source could be a different API, or any other publicly available information (for example, a .csv file)
- The database should be populated with at least 10,000 records and have at least 5 tables.
- Your application should use at least 7 different SELECT queries:
    - One full-text query (make sure the data is indexed correctly)
    - 6 complex queries (e.g. nested queries, group by, aggregations, EXISTS, etc.)

Project requirements:

- Your application should **not rely on continued usage** (for example, interactions with different users).
- The MySQL server you should use is "mysqlsrv1.cs.tau.ac.il". Users and password will be given after you sent me an email with your group details (full names and ids)

Working Teams:

- The work should be done in pairs (could be the same as HW1 submission).
- After I register your group, you will be given a user for mysqlsrv1.cs.tau.ac.il
    - All DB interactions should be done with this user

Coding Guidelines:

- The creation of the DB is eighter in sql or in python.
- It's update should be written in python.
- The different SQL queries that your application produce should also be in a python file.
- You can use external libraries, if:
    - It does not automatically create a DB schema
    - It does not perform any database optimizations
    - It does not generate SQL queries automatically

- Your code should be readable and documented
- Errors should be handled

DB Design:

- The DB should be designed according to the principals taught in class, including:
  - Meaningful names for tables, columns, indices, keys etc.
  - Primary and foreign keys
  - Use of indices to optimize your queries
- Explain why you chose your DB design:
  - Why do you think it is the most efficient for your needs
  - What are the disadvantages of other designs you have considered

Documentation:

- User manual:
  - What does the application do? (overview)
  - The design of the application: How it would have looked like if it was implemented (you can design it in power point or word document). What are their features?
- Software documentation:
  - DB scheme structure (also explain your choices while designing the DB)
  - DB optimization performed (e.g. your use of indices)
  - Description of your 7 main queries:
    - What does it do
    - How did you optimize this query
    - How does your DB design supports this query
- Code structure
- Description of your API + how did you use it
- General flow of the application

Submission:

- The MySQL server should be “mysqlsrv1.cs.tau.ac.il” with the user you will be assigned
- You should submit your source code and the documentation in a single .zip file in the following structure:
  - /SRC
    - /API-DATA-RETRIEVE (the code which inserts the data to your DB)
    - /CREATE-DB-SCRIPT (an .sql or .py script which creates the DB)
  - /DOCUMENTATION
    - /NAMES-AND-IDS.txt (group members details)
    - /USER-MANUAL.pdf (see “documentation” section)
    - /SOFTWARE-DOCS.pdf (see “documentation” section)
    - /MYSQL-USER-AND-PASSWORD.txt (the MySQL user you were assigned)

Tips and advices:

- Make sure you invest efforts in the right places. This is a Database project, and the focus should be on the database design, optimization, and queries rather than the UI
- All APIs have a daily/hourly usage limit. Make sure you retrieve data efficiently so you're not exceeding the limit. Make sure to start this process in advance as it might take several days.
- For any system issue please contact system@cs.tau.ac.il and address the course staff if your issues are not resolved

Good Luck!