

CME2202 Data Organization and Management

Spring 2023

Assignment 2

Deadline: 14 May 2023, 22:00.

Problem: Design a simple database query system. The system must accept a CSV file (students.csv) as input as the initial database. After initial loading of the file, the records must be sorted by their index. The query will be given in SQL with a restricted command set for simplicity (see below for details). The resulting data set (result of the query) must be stored as a JSON file.

Methods:

- You must use the Python programming language and your code must be in a single PY file. We recommend Anaconda Navigator package with Spyder IDE.
 - You can use another IDE (like PyCharm or VS Code), but we will not help with problems arising from the selection of a different IDE.
 - You can use Jupyter Notebook (Also in Anaconda package) to test and debug your code, BUT your final submission upload must be in stand-alone Python code with PY extension.
 - File name format: "GroupNo_Student1 Number_Student2 Number.py"
- You can use any Python library shown in the course, libraries for reading CSV and writing JSON files. However, you must:
 - Install the libraries yourself by following the installation procedures in the webpages of the libraries.
 - NOT use a library for query parsing.
 - If you are unsure of whether you can use a new library you find, write a post in the Forum and get our confirmation.
- Accept the query from the user and give an error message if the query doesn't fit the format of simplified SQL given below.

Simplified SQL: It has the following form:

```
SELECT {ALL|column_name} FROM STUDENTS WHERE  
{column_name|=,!=,<,>,<=,>=,!<,!>,AND,OR3} ORDER BY {ASC|DSC}
```

```
INSERT INTO STUDENT VALUES(val1,val2,val3,...)4
```

³ There can be at most two conditions bound by an AND or OR keyword. So we do NOT have cond1 AND cond2 OR cond3. We can only have cond1 AND cond2 or cond1 OR cond2.

⁴ All values for the new student record must be given, if incorrect number of attributes are entered, you can raise error.

DELETE FROM STUDENT WHERE {column_name|=,!=,<,>,<=,>=,!<,!>,AND,OR }⁵
exit

Examples:

SELECT name,lastname⁶ FROM STUDENTS WHERE grade != 40 ORDER BY
ASC

SELECT name FROM STUDENTS WHERE grade > 40 AND name = 'John'
ORDER BY DSC

INSERT INTO STUDENT VALUES(15000,Ali,Veli,ali.veli@spacex.com,20)

DELETE FROM STUDENT WHERE name = 'John' and grade <= 20

Note: For string values like name, the condition operators are limited to {=, !=}.

- After finishing the query, the program must wait for another query, or quits if the user enters exit.
- You can assume queries as case insensitive.
- If a column or table is dropped, and used in

a later query, you can print an error message.

Rules:

- Don't wait until the last week/day!
- The deadline is final, for late uploads points will be deducted.
- Your groups at Assignment 1 are valid for this project as well. You can't change groups after this point, but if an unforeseeable event (like sickness) changes the group, please notify us immediately.
- The assignment will be uploaded by the first member of each group.
- Your code must include comments where needed, explaining how your code works and your reasons. Your student numbers and names must be comments at the top of your code.
- All questions about the assignment must be addressed by e-mail at atakan@cs.deu.edu.tr. Please avoid questions like "my code doesn't work; can you find the problem?".
- Please state your student number in all correspondence.

Good Luck.

⁵ There can be at most two conditions bound by an AND or OR keyword. So we do NOT have cond1 AND cond2 OR cond3. We can only have cond1 AND cond2 or cond1 OR cond2.

⁶ In the query submitted by the user, there will be no space between column names, if there is you can raise an error.

⁵ There can be at most two conditions bound by an AND or OR keyword. So we do NOT have cond1 AND cond2 OR cond3. We can only have cond1 AND cond2 or cond1 OR cond2.

⁶ In the query submitted by the user, there will be no space between column names, if there is you can raise an error.