

CS430 Homework 3

Brendan Nguyen

Due: Wednesday, Mar. 15, 2023

Question 1 (40 points)

Given the following DB schema:

- *Customers* (cid: int, name: string, city: string, state: string, age: int)
- *Has_account* (cid: int, aid: int, since: date)
- *Accounts* (aid: int, atype: string, amount: real)

Primary keys are uid in each relation. Relation *Customers* contains information about customers. A customer is uniquely identified by *cid*. Relation *Accounts* contains information about bank accounts. An account is uniquely identified by *aid*. The type of the account is given in column *atype* (for example *atype* = 'savings'). *Customers* have *Accounts*. For a customer that has an account (i.e. is the owner of that account) there is a record in relation *Has_account*, that has the *cid* of that customer and the *aid* of that account, as well as the date when the account was opened (column *since*).

Using this DB schema:

- Write the SQL statements to create the tables from this DB schema. Do not forget to define the necessary key constraints. The statements should be written in an order such that if executed in that order will not cause an error.
- Write SQL statements to insert two records in each table. The statements should be written in such an order such that if executed in that order do not cause an error.
- Write the SQL query that extracts the id, name of customers from city "Boston" that have some account with an amount smaller than 5000. The result should not contain duplicates. Sort the result by name in descending order.
- Write the SQL query to extract the id, name, and age of customers who did not open any account between Jan 1st 2020 and Dec 1st 2021 (including these dates).
- Write the SQL query to extract the id, name, and age of customers who have both "savings" and "checking" types of accounts.
- Write the SQL query to extract the name and id of customers who have at most 10,000 across all their accounts.
- Write the SQL query to extract all account ids of type checking that have at least 2 owners.
- Write the SQL to get the number of accounts each customer older than 25 and younger than 35 has. In the result include only customers who have at least 2 accounts.
- Write the SQL to extract the id, name, and age of customers who opened accounts both in year 2018 and 2020.
- Write the SQL query to extract the id and the name of the customers who are from "MA" state and who have at least 2 accounts of type "savings".

Question 2 (60 points)

Given the following DB schema:

- *Articles* (aid: int, aname: string, first_author: string, pubyear: int, pubcompany: string)
- *Students* (sid: int, sname: string, age: real, state: string)
- *Reads* (sid: int, aid: int, year: int)

Primary keys are underlined in each relation. An article is uniquely identified by `aid`. An article has an id (`aid`), a name (`aname`), one first author (`first_author`), a publication year (`pubyear`), and a publishing company (`pubcompany`). A student is uniquely identified by `sid`. A student has an id (`sid`), a name (attr. `sname`), age (attr. `age`) and a state (attr. `state`). If a student reads an article, a record will be present in the *Reads* relation, with that `sid`, `aid`, and the year the article was read.

For this DB schema:

- Write the SQL statements to create the tables from this DB schema. Do not forget to define the key constraints. The statements should be written in an order such that if executed in that order will not cause an error.
- For each table, write an insert statement to insert one record. The statements should be written in an order such that if executed in that order will not cause an error.
- Write the SQL statement to find the number of articles that have a `first_author` whose name contain string "an". The query has to be case insensitive with regards string "an".
- Write the SQL statement to find information about the newest articles (hint: `pubyear` is max). Sort the result by the name of the article in ascending order.
- Write the SQL statement to find the id, name and age of students who read all articles. The result should contain no duplicates.
- Write the SQL statement to find the id and name of students who read some articles published in 2020 and did not read any article published in 2018. The result should be sorted by the name of the students in descending order.
- Write the SQL statement to find the ids and names of the students who read some articles in the same year when that article was published.
- Write the SQL that the number of articles read by each student. The result should contain information only about those students who read at least 3 articles.
- Write the SQL statement that extracts the minimum, maximum and average age of students for each state. Show this information only for states where there are at least 2 students.
- Write the SQL statement to find the articles whose first author name (`first_author`) starts with letter "B" and that are published either before 2018 or after 2020 (note that 2018 and 2020 are included).
- Write the SQL statement to find the id, name of the students who read all articles published in year 2022 by the publishing company "penguin". The result should contain no duplicates.
- Find the id, name, age, and state of the students who did not read all articles published by "simon" publishing company. The result should contain no duplicates.