# CS630 - Homework 3 100 points

#### INSTRUCTIONS

You must copy your solution to your Unix account, according to the following instructions.

- Answers to question 1 and question 2 must be written and saved in files named Q1.sql and Q2.sql respectively and copied under a folder named hw3 in your cs630 folder. Instructions on how to create the hw3 folder is the same as in assignment 2. Please ensure that the permissions on the files and folder are set correctly, as per instructions (same as in assignment 2).
- You are allowed to copy your solution files and make changes to them any number of times before the due date, but no copying or modification is allowed after the due date. Copying a new version or modification of either file will invalidate your submission, which causes no credit to be given to your work.
- Please ensure you have access to gradescope, Unix account and Oracle before the due date. Unlike the second assignment where exceptions were made, this time no exceptions will be made for lack of, or issues related to access.

#### **Important Notes:**

- SQL statements must run against the Oracle database we use in class. Test your queries
  against the Oracle DB with the necessary data/records to ensure the correctness of your
  solutions. SQL queries that do not run successfully against the Oracle DB will receive no
  credit.
- Do not use SQL constructs that we have not covered in class.
- Each SQL statement should end with a semicolon (;).
- In Q1.sql and Q2.sql files, before each SQL statement you MUST include a comment line with the problem number the sql statement is for (e.g., before writing the SQL query for (c) add a comment line such as --Answer for c). Remember that a comment line starts with two dash symbols. Feel free include additional comments if you like.

### Question 1 - 40 points

Consider the following DB schema.

Primary keys are underlined in each relation. Customers relation contains information about customers. Customers are uniquely identified by their ids (cid). Accounts relation contains information about bank accounts. Accounts are uniquely identified by their ids (aid). The type of the account is given in column atype. Customers have accounts.

Ownership of accounts by customers is reflected in has\_account relation. For instance, customer 123 owns account 456. The since filed stores the date when the account was opened.

Using the above schema, write the necessary SQL statements to do the following. Each question has 4 points.

Note that you may need multiple statements for some of the items. Further note that the order of the statements "may" be significant in some cases. In such cases, you must ensure that your statements are written and save in the solution file in the correct order. If the statements are executed in the order they are written successfully, it means they are in the correct order.

- a) Create the three tables in the schema. Do not forget to define the necessary key constraints.
- b) Extract the id, first name and last name of customers from "MA" state, who have at least one account with an amount greater than 1000. The result should not contain duplicates and should be sorted by the last name in an ascending order.
- c) Find the number of accounts for each type of account in the database.
- d) For each customer from "MA" state has who has at least 2 accounts, show their id, first and last names, along with the total number of accounts they hold.
- e) Extract the id and the first name of the customers who have at least 2 accounts of type 'checking'.
- f) Extract the id, first name, and last name of customers who opened at least one account in year 2018 AND in year 2020.
- g) Extract the id and the last name of customers who did not opened any account after Jan 1<sup>st</sup>, 2020.
- h) Extract the id and the last name of customers who at least one "saving" and at least one "checking" account.
- i) Extract the id and the last name of customers who have at least 20,000 across all their accounts. That is, the sum of the amounts in all their account is 20,000 or more.
- j) Extract all account ids of type checking that have at least 2 owners. Note that the schema allows a single account to be owned by more than one customer.

## Question 2 - 60 points

Consider the following DB schema:

Primary keys are underlined in each relation. A book is uniquely identified by bid. In addition to id, every book has a name (bname), an author (author), a publication year (pubyear), and a publishing company (pubcompany). A student is uniquely identified by sid. Additionally, every student has a name (sname), an age (age), and a state (state). If a student reads a book, a record will be present in the Reads relation, with the id of the student and the id of the book that was read, along with year in which the reading took place.

Using the above schema, write a statement, or statements to do the following. As stated before, for multi-statement answer are written and saved in the correct order. Each question has 5 points.

- a) Create the tables the three tables defined for the schema, in the correct order. Do not forget to define the key constraints.
- b) For each table, write an insert statement to insert one record (order matters).
- c) Find the students who live in 'MA' and they are either younger than 25 or older than 35.
- d) Find the number of books whose author's name starts with a **B** or a **b**.
- e) Find the oldest books in the database (hint: pubyear is min).
- f) For every state that has 50 or more students, extract the average age of all students in that state.
- g) Find the id and the name of students who read all books in the books table. The result should contain no duplicates.
- h) Find the id, the name and state of the students who read all books published by 'penguin' publishing company. The result should contain no duplicates and sorted by the name of the students in descending order.
- i) Find the id and the name of students who read any book published by 'penguin' and read no books published by 'simon' (penguin and simon are publishing companies).
- j) Find the names of all students who read at least one book in the same year as the book was published.
- k) Find the name, age and state of the students who did not read all books published by 'penguin' in year 2020. The result should contain no duplicates.
- I) Extract the id, the name, and the author of books published by 'penguin' and were read by all students. The result should contain no duplicates.