CSCI 4125/5125 Course Project
Data Models and Database Systems
Spring 2022
Course Project
Phase 3: Introductory SQL

Due: Tuesday, 3/22 @ 11:59pm
Reading: Silberschatz Chapters 3.1 – 3.4, 3.6 – 3.9, 4.3 – 4.5

Submission Guidelines:
1. This assignment is worth 100 points for all students.

2. It is your responsibility to make sure all files are readable and submitted on time.
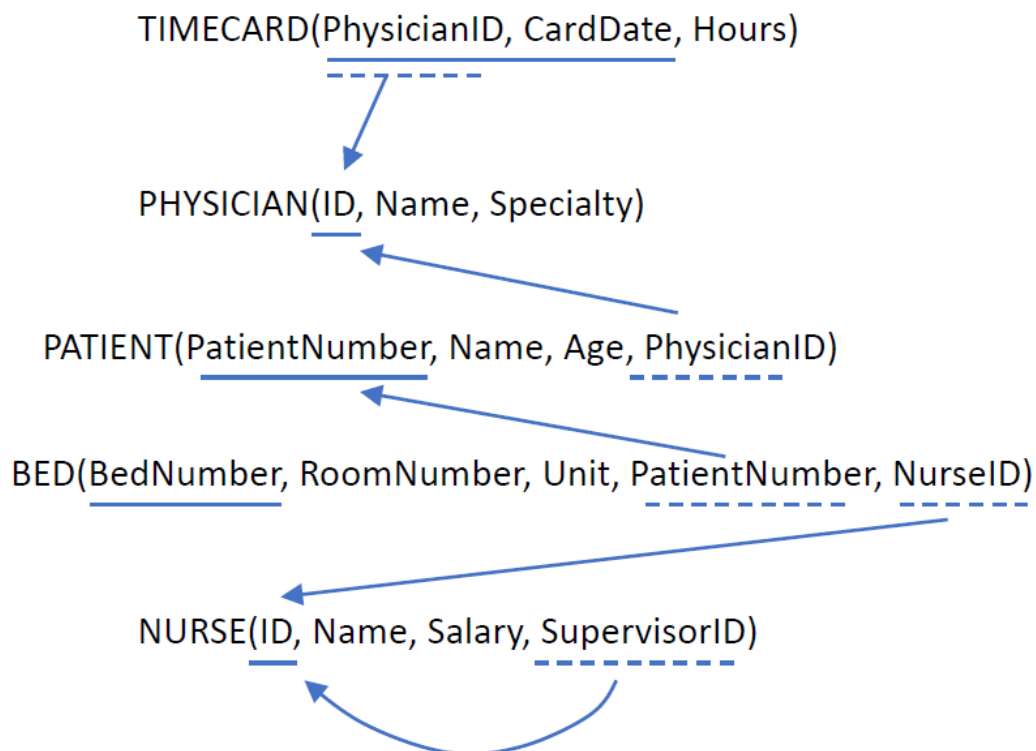
Submission:
- Part A requires you to submit a single .sql file worth 30 points.
- Part B requires you submit five .sql files and another file in the format of your choice worth 30 points.
- Part C requires you to submit a single .sql file and a single .txt file worth 40 points.

Part A. SQL DDL (30 points)

Write a SQL DDL script to create your tables from Phase 2. Use the following guidelines and relational schema:

- At the top of your script include a drop table command for each of your tables. Note that you must pay attention to referential integrity when considering the order to drop tables.
- Columns names should match the attributes in the relational schema. This will allow for consistency in our queries later.
- Columns must use reasonable domains based on the data in the included .csv files.
- All primary keys must be declared.
- All foreign keys must be declared.
- Run your SQL script and debug any errors.
- **Submit:** a single .sql file named hospital_schema.sql.

TIMECARD(PhysicianID, CardDate, Hours)

PHYSICIAN(ID, Name, Specialty)

PATIENT(PatientNumber, Name, Age, PhysicianID)

BED(BedNumber, RoomNumber, Unit, PatientNumber, NurseID)

NURSE(ID, Name, Salary, SupervisorID)

Part B. Populating the database (30 points)

Using your Java program from Phase 2, generate INSERT statements for the _five_ .csv files
included with this document. You should generate a total of _five_ SQL scripts containing INSERT
statements. If are not able to get your Java program working, you can create the INSERT
statements manually, but note that this will be tedious. Use the following guidelines to submit
your work:

- Data types must be properly formatted, e.g., strings must use single quotes, dates must
  use the correct format.
- Each script should include a commit (i.e., "commit;") at the end of the file. You can
  modify your Java program to simply write that before you close the outfile.
- Name each file [table name].sql.
- Run each script in SQL Developer. Remember that you can run SQL scripts using
  @[path]\[file].sql
- You will need to run the INSERT's in the proper order that does not violate referential
  integrity.
- Read the output that is generated. If there are any errors, it is up to you to fix them.
  Errors to watch out for include, improper syntax, violating domain constraints, violating
  primary key constraints, and violating referential integrity.
- **Submit:** Five .sql files containing your insert statements. Name them [table].sql. Also
  submit in any file format (e.g., txt, doc, pdf, etc) you want that explains the order in
  which you ran your scripts (that doesn't violate referential integrity).

Part C. SQL Retrieval Queries (40 points, 5 points/query)
Write queries for your respective database. **Submit**: your answers in a single .sql file named hospital_queries.sql and the outputs from each of your queries in a .txt file named hospital_results.txt.

Mountain Valley Community Hospital

1. Find all nurses (their names) with a salary greater than $85,000.

2. Find all the "surgery" (for the specialty) physicians and sort the query output by names (any direction).

3. Find all physicians with medicine in their specialty name (hint: remember the LIKE operator and case sensitivity).

4. Find all nurses who do not have a supervisor.

5. Find the names of all nurses with a salary between $70,000 and $80,000 (inclusive).

6. Find the names of the physicians who have a specialty containing "ology".

7. Find the minimum and maximum salaries amongst all nurses. Use only one query.

8. Find the average salary for all nurses.

I am also including two more challenging queries that you may attempt to test your knowledge of SQL. They aren't mandatory – you will get 100% credit for the homework if you correctly answer queries 1-8. (+2 points each)

9. Find the name of the nurse that has the highest salary. Do not hardcode any salaries or other values – you must use SQL without assuming you know the current database snapshot.

10. Find the nurses with a salary less than the average overall salary for all nurses + 20% (i.e., less than 1.2 * average salary). Do not hardcode any salaries or other values – you must use SQL without assuming you know the current database snapshot.