CSE 4020/5260 Database Systems

Spring 2023

Final Project

Total Points: 25

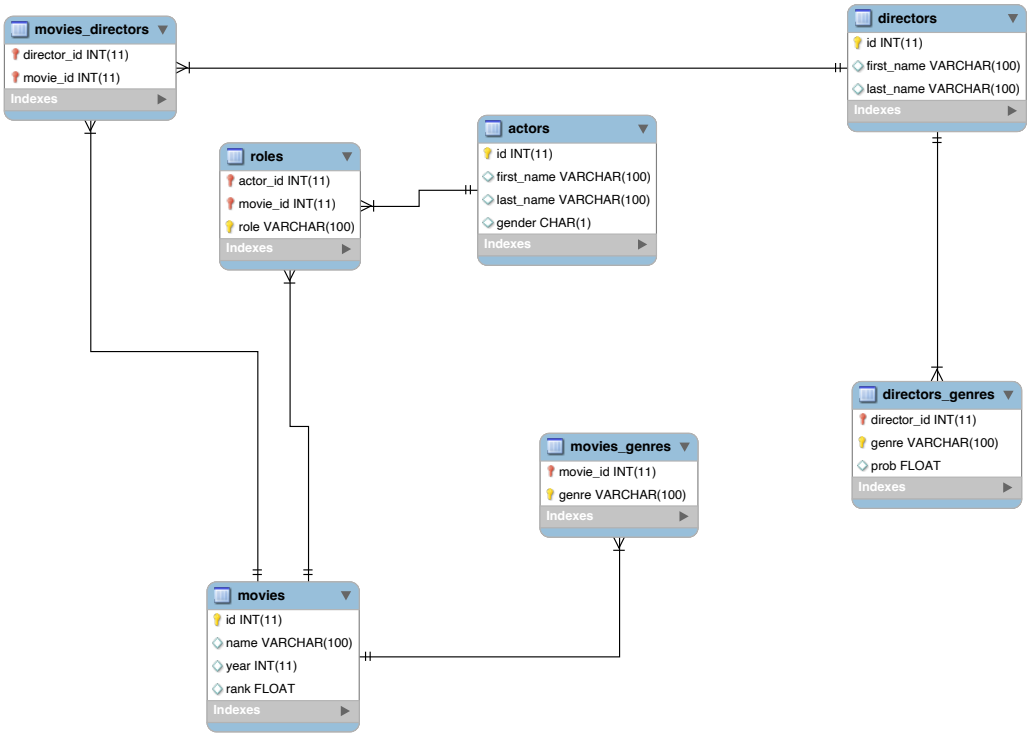
Date Assigned: Friday, Apr 14, 2023 **Due Date:** Monday, Apr 24, 2023

Instructions: Submit your work on Canvas as a Jupyter Notebook `ipynb` file named `cse4020_yourname_final_project.ipynb`. Make sure to add headings and question numbers to your notebook using Markdown. Include in your notebook the contents of your DDL file, your code, and the corresponding output.
Important: You must complete this assignment using code in an imperative language such as Python, Java, C, or any other imperative language that works in your Jupyter notebook. Simply executing queries in a GUI will result in a grade of 0. You may use the Java and Python code provided by the instructor to batch execute a DDL file as well as other example code demonstrated in class.

Key Concepts Demonstrated

- Executing DDL & DML Statements on a database
  - Reducing an ER Diagram into tables
  - Creating a DDL File based on a set of tables
  - Coding in a Jupyter notebook
  - Writing program code that reads a DDL file and executes the statements
  - Issuing DDL statements on a SQL database
  - Issuing DML statements on a SQL database

1. (5 points) Create a DDL file based on the following ER diagram (You may use dbdiagram.io as demonstrated in class or you may create the tables manually). Make sure to represent the keys (primary and foreign) in your DDL. Add your DDL file as Markdown color-coded SQL statements to your notebook.



2. (5 points) In a Jupyter Notebook, write code either in Python, Java or C, or another imperative language, to create a database named imdb_movies; then use code to connect to your DB instance and create each of the tables in your DDL file. Do not rename the tables or any attributes in the ER diagram.
3. (5 points) Write code using the language chosen in question 2 above to connect to your DB instance and insert the data from each of the DDL files provided on Canvas into the your imdb_movies database. Display the number of rows affected for each table (1 cumulative count for all inserts per table - example: 388269 rows inserted for table movies).
4. (10 points) Write and execute a query that returns a count of all movies in the database where the *rank* column is not NULL. Print out only the count.