CSE 412: Database Management
Spring 2017
Group Project – Phase 2
Due: March 12th, 2017, 11.59pm

The group project constitutes 25% of the total grade for this course, you can obtain at most 10 points from Phase 2.

- In this phase, each group will have to perform the following tasks:
  - **ER-to-Relational:** Transform the ER diagram into the relational model using SQL Data Definition Language (DDL). Then, you have to create the database that follows that relational model by using a DBMS of choice. Justify in the report the choices you make when implementing a translation strategy from ER to SQL.
  - **Sample Data:** Fill in the database with data synthetically generated, i.e., you need to write scripts to generate data suitable for your database. Data can be directly inserted in the DB or created as CSV values and then loaded.
    - If you fail to write a script to generate data, another option is to have at least about 100 manually created tuples to verify your queries. However, it is required to generate at least 1000s of tuples to get the full grade.
    - If you provide a script to generate data, write a short ReadMe explaining how to run it.
  - **SQL Queries:** This is the most important task in phase 2. Prepare the SQL queries that cover the application description detailed in phase 1.

More details on the DB
- It should have at least 8 tables, they must have primary keys.
- Among all tables, there should be at least 6 foreign keys; specify their behavior when there is a DELETE or an UPDATE.
- Make sure to have at least one check constraint in each table definition and at least 2 assertions involving two tables.
- Remember to provide the DDL scripts to create your DB

More details on the queries
- Include at least 8 queries. Among these there will be at least:
  - Two queries with JOIN conditions.
  - Two queries with GROUP BY and aggregation.
  - One correlated nested query.
  - One query to insert/update the data with user interaction
    - It can be a simple insert of a new tuple.
    - This is required even if you are loading the data with bulk loading from text file and your application will not be focus on content provided by users.

Optional: add at least one index. It would be nice to show performance differences with/without index for an expensive query (you need more than 100 tuples for this task).

In the report it is not necessary to copy all the details that are in the DML SQL scripts, as these scripts are submitted as separate files in your submission. The submitted SQL scripts will be used during your demo, so **make sure they are correctly running on your data**. Time is going to be constrained for the demo, so make sure to use it wisely to show all the great features of your queries.

Guidelines and recommended structure for the report
- One page overview of the application with list of main functionalities.
- One page with the ER diagram and notes/comments/annotations on it (highlight changes from project phase 1, if any).
- One to two pages to describe how the ER diagram map to the create statements, no need to copy the entire SQL code here. Describe check constraints and assertions.
- One to three pages to comment on your SQL queries: how they support the functionalities, no need to copy the entire SQL code here.
- Make sure tables and constraints are consistent with the ER diagram for your DB, and vice versa.
- List any FD that is not a primary key and clarify if it is enforced in your DB.
- When writing the queries, use indentation and capital letters for the reserved keywords.

Each group will therefore submit a **report, the DDL and DML scrips, and instructions to run the data generator (if any, otherwise the dataset),** as well as demonstrate the aforementioned tasks to the instructor by using a **SQL console**. This phase does not require any GUI demonstration.

Please **upload your submission in Blackboard** as a compressed file ZIP file with the following name format: "SecondPhase_GroupID.zip", with ID being your group number (e.g., "SecondPhase_Group10").

Notice that you are required to submit also the **individual peer evaluation** (the link to the online form will be posted in Blackboard). Your contribution to the group work is part of the grade that falls under "Participation" (10% of the total).