# CS482/502 Database Management Systems I (Spring 2020)

# Project: Phase 2 (80 points)

**Requirements**

The purpose of this project phase is to create a web interface for inserting large amounts of data into the MySQL database that was created in Phase 1.

**Tasks:**

- Create an interface that interacts with the database created in Phase 1. The interface should be able to execute the following functions:
  - Users should be able to insert a text file that contains the data for the corresponding table. E.g. If the name of the text file is *Players.txt*, then the data in the text file should be inserted into the *Players* table. The format of the textfile will be: *Column _value1*, *Column value2*,..., , *Column valueN* where a table consists of *N* columns. The text file will not contain any headers.
  - Before insertion of the data, the user should be given 2 options to choose from: *Single Insertion* or *Bulk Loading*
  - These techniques are explained in more detail in the following links:

    *Single Insertion* : https://dev.mysql.com/doc/refman/8.0/en/insert.html

    *Bulk Loading*: https://dev.mysql.com/doc/refman/8.0/en/load-data.html

  - Implement these two different ways of inserting data: *Single Insertion* and *Bulk Loading.*
  - Implement an option to delete all data from a particular table. Remember, we will not be changing the schema of the database.
  - Make sure that the necessary constraints **are enforced** while inserting and deleting the data. Your system should output an error message if any constraints are violated during insertion or deletion of the data.
  - There should be a textbox to query the tables (and the database) that were created in Phase 1.
  - Users should be able to write a functional SQL query. If the query is correct, then your system should display the correct results. If there are errors, the system should display the appropriate errors.
  - Along with the results, report the time it took to execute the query.

- A detailed report should be submitted. The report should contain the following:

- Description of the methods/techniques used to complete the tasks.
- An analysis (using charts) on the performance of *Single Insertion* vs *Bulk Loading*.
- You will need to create 3 datasets: the size of the Players table should be 100K, 150K, and 200K respectively for these 3 datasets. Report the times it took to execute *Single Insertion* and *Bulk Loading* for these 3 datasets. Create appropriate charts to show your analysis.

- After the project has been submitted, you need to record a video which should include the tasks listed below. You can use applications like BandiCam or OBS Studio for recording in Windows. If you are using Linux, it has a default screen recorder.

- Tasks (in order) to include in your video:

1. Insertion of dataset 1 using *Single Insertion*. The time should be calculated and shown in the video.
2. Deletion of dataset 1.
3. Insertion of dataset 1 using *Bulk Loading*. The time should be calculated and shown in the video.
4. Deletion of dataset 1.
5. Prove that your *check constraint* works while inserting the data (insert data that violates the check constraint and show the error). The video should also show what was *wrong* with the data that you tried to insert.

**Handing in & Grading:**

- Combine and compress all source files and report to a .zip file with name "FirstNameLastNamePhase2.zip". See course syllabus for policies on late submission and plagiarism.