CSCI 585, Summer 2015, Assignment 1

Due dates: Part I: June 19; Part II: June 26.

1 ER data model

Design a schema that incorporates the specification described below as efficiently as possible. If something is not clear, you can ask on D2L or make any *reasonable* assumption, as long as that does not contradict with what has clearly been said in the assignment or assumed by you elsewhere.

1.1 Design Specification

Design the database system for a website about golf. It should store and manage the following information. In case of ambiguous parts, make a valid assumption about how you would expect a real system to work; if you aren't sure, please ask. Please write any assumptions you make on your submitted work.

1.1.1 Courses

A course has a name; no two courses have the same name, and every course has a name. Each course has an associated *GreenFee*, the amount of money (charges can be dollars and cents, but may not be fractions of cents) it costs to play that course.

1.1.2 Tees

Each tee is associated with a course in §1.1.1, although any given Course may have multiple Tees. Each Tee has a distinct name *among those associated with that Course*.

Each Tee has associated with it a real number (two digits, plus one after a decimal point) called the CourseRating, an integer SlopeRating, and an integer yardage.

Clarifying Information: The following should help you understand the requirements of his section. When one plays golf, he or she goes to a Course, and selects a set of Tees to play from. These are physical markers on the course that denote where one begins or resumes play at varying points in the round (for more on rounds, see § 1.1.4). You can think of the various Tees as "difficulty settings" for the Course: a novice might play from the easiest set, an expert might play from the hardest, and so on. The CourseRating and SlopeRating are among the ways for golfers to figure out just how difficult a particular Tee is.

1.1.3 Golfers

Each Golfer is identified by a distinct GolferID, although he or she may also have a name. Each golfer also has a "home course", which must be a course in §1.1.1.

While you don't need this fact for the schema, you should know (perhaps for one or more queries) that a golfer is said to be a "member" at his or her home course.

1.1.4 Rounds

When a Golfer plays golf, this is known as a "round" of golf. A round of golf is played by a Golfer, specified by their ID from §1.1.3. Each round of golf has one particular day on which it was played. For purposes of assignment, input will contain no more than one round per day per golfer. Students do not need to enforce this.

Each round played occurs at a Course (§1.1.1), is played from a set of Tees (§1.1.2), and has an associated integer score.

While you don't need this fact for the schema, in golf, a *lower* score is better.

1.2 Submission Guidelines

By 11:55 P.M. (Pacific time) on June 19, 2015, submit a **PDF document** to D2L containing your data model. Late days may be used for this as usual, although they *do not* count "double" – that is, using a late day for assignment one, part one does not also allow you to submit the full assignment one a day late "for free."

2 Map the ER model into MySQL

Install a MySQL workbench for your computer so that you will be able to create the required files for the submission. You will be mapping your ER model into MySQL by creating tables and importing data. The data will be available on D2L by June 19 in a reasonable format. You will also be creating queries to answer questions about the data.

The queries below will require that every part of the Schema described in the Design Specification be mapped into a MySQL table.

2.1 Guidelines

Use reference for **foreign key** as appropriate.

Do not use triggers.

3 Simple queries on the database

Write the following queries in SQL and run them on your database as developed in the previous part. Depending on the data, your query might not return any data but that does not mean your query is wrong.

In queries where you are asked to "find golfers," returning the GolferID is sufficient unless the problem statement says otherwise.

1. How many distinct golfers are in the database?
2. List all courses that have one or more tees that play more than 6680 yards.
3. Which golfers are members at Wilson?
4. How much did Logan spend on golf in 2013?
5. Give a list of every course that 'Michael' has played. Do not list duplicates.
6. How many different golfers have played 'Harding' from the Blue tees?
7. List the ten best (lowest) scores during 2014 and the golfers who produced the score. If one golfer appears multiple times, he or she can count multiple times. If there are more than ten scores in the "ten best," then as long as you list ten scores such that no unlisted score is strictly better than any listed score, you're good.
8. What was the worst score recorded from the Yellow tees at Rancho in the timeframe covered?
9. Find the average score at Rancho between June 1 2013 and August 31 2013.
10. What is the average score of members at 'Wilson'?

3.1 Submission Guidelines

1. Your submission of part2 and part3 should include one `createdb.sql` file, one `dropdb.sql` file, one `projone.sql` file, and a `readme.txt` file. Your readme.txt file should include your name and student ID and anything else you think we need to know about your submission.
2. createdb.sql file should create required types, tables, primary keys, foreign keys, and populate all data provided. **If you do not do this step, we cannot evaluate your queries or your ability to clean-up the database, and those will receive automatic zeroes.**
3. The dropdb.sql file should drop all types and tables that are created by createdb.sql. There is a 10 point penalty if this file is missing from your submission or if it does not drop all of your database objects.
4. The `projone.sql` file should contain the ten queries, in the order they're listed. Include a comment before each query indicating the question number. If you omit one, list that as a comment in place of the query.