# **CSCI X370: Database Management**

**Spring 2018**

## **Project 3: Performance Tuning of SQL Queries**

## Due: Feb 26, Monday (11:59 pm)

In this project, you need to test MySQL using the queries given below.

Consider the below schema and tables:

1. Student [ id, name, address, status ]
2. Professor [ id, name, deptId ]
3. Course [ crsCode, deptId, crsName, descry ]
4. Transcript [ studId, crsCode, semester, grade ]
5. Teaching [ profId, crsCode, semester ]

Tuple counts: Student: 10,000; Professor: 1,000; Course: 2,000; Teaching: 5,000; Transcript: 5,000

You can assume any reasonable datatypes or you can refer to the tables given in Chapter 3 of textbook.

Write queries on the given schema as below:

1. List the name of the student with id equal to v1 (id).
2. List the names of students with id in the range of v2 (id) to v3 (inclusive).
3. List the names of students who have taken course v4 (crsCode).
4. List the names of students who have taken a course taught by professor v5 (name).
5. List the names of students who have taken a course from department v6 (deptId), but not v7.
6. List the names of students who have taken all courses offered by department v8 (deptId).

In these queries, v1, v2, ... stand for value1, value2, ...etc.

To be able to test the MYSQL database system, you need to analyze query plans generated by the DBMS and tune these queries to be able to run them faster. For tuning these queries, a link to tuning guide for MySQL is given below:

http://dev.mysql.com/doc/refman/5.7/en/optimization.html

You will need to run the queries, look at the query plans, and use some of hints given in the tuning guides (removal of redundant parenthesis in the query, rewriting queries etc.). Also, you can add indexes to try to speed it up. Then you need to write a brief report about each query, which indicates what you did to speed it up. For doing this, you need to obtain initial timing result and what the timing result is after you tuned it.

First of all, you need to create a database, which you can run those queries on. You'll create a sample database used in the textbook, and you can find the details of this database in the book (database on students, courses etc.). However, if you just create a database with small number of tuples it is going to take 0 times, and you will not be able to make any comparison. Hence you need to insert enough tuples as per the given tuple counts to get the proper timing results and can make a comparison in between. You need to use Tuple Generator to be able to quickly populate your database.

You need to turn in before and after .sql files and query plans for six queries (see above). You can do this using MySQL workbench. Also, please analyze the queries using explain plan and write your observations and performance cost and time taken to execute before and after the tuning.

**What to submit:**

The following documents should be submitted:

1. Schema file (TeamName\_DatabaseName\_MySQL.sql)
2. Queries before tuning (TeamName\_BefTuning\_MySQL.sql)
3. Queries after tuning (TeamName\_AftTuningMySQL.sql)
4. Document with explain plans for each of the queries (i.e., the screen shots of the explain plans for each of the queries) and your observations about how the performance has improved after tuning the queries.
5. Readme file

**Note 1:** The .sql files should include the create database and use database statements and comments added in front of them.

Please add comments in the database file before each query related to the question number and query.

e.g -- create database TeamName

-- use database TeamName

**Note 2:** The explanation for each query, before and after tuning should be in both graphical as well as description in words. Also, mention clearly the speed up between the corresponding tuned and un-tuned queries.

You can generate the query plan directly from any workbench you use for query execution.

**Note 3:** Please add a class file (into Tuple Generator) that connects to the database and write JDBC queries for storing the results.

**How to Submit:**

Submit your ".zip" file using ELC. Only team leaders need to make a submission. **Every student needs to submit a peer-evaluation form within 24 hours of the project submission deadline**.

Do not place your solution on a public web site. Submit your own work and follow the course misconduct policy.