ASSIGNMENT 1

10 %

Use Create, Insert, Update and Select statements to store, and retrieve, data in a database. The assignment must be done in groups.

Lyle Wood
Lyle.Wood@GeorgianCollege.ca

COMP 2003 Relational databases

Assignment 1 – Relational Databases

This assignment requires you to create, populate, manipulate, and retrieve data using SQL commands. Your mark in this assignment counts for 10% of your final grade.

Group Allocation

This is a group assignment. Please enroll in one of the groups on the course website. It is your responsibility to enroll with a group. Each group has a maximum number of members as detailed on the course website.

To facilitate your collaboration, a private discussion board and online file storage has been provided for each group. Feel free to use other online collaboration tools at your will, provided all group members agree to that tool's use.

Assignment Requirements:

Please submit your work as a WORD document – docx (not pdf) – as well as your finalized SQL script (.sql).

Report formatting

- Name your report file *GroupName*_Assignment1.docx
- Give your report an appropriate title and include a title page.
- Please include the name and student number of all participating group members
- Before providing an answer, please include the question's number and text.
- Please include the text of your queries (including comments) and a screenshot which clearly shows the execution of your commands and the results. You might need to send more than one screenshot to show the results for each question.

SQL formatting

- Name your SQL script *GroupName*_Assignment1.sql
- Please include the name and student number of all participating group members
- Complete all of your SQL work in *GroupName*_Assignment1.sql
  - Inside I should find all SQL commands required to complete the assignment.
- The data provided (barrie_weather.sql) is solely for the purpose of populating your database with data – do not complete your SQL work in this file.
- Utilize the SQL Standards document provided in week 2 – with note to capitalization of keywords and lower-case database, table and field names.
- Remember to add a brief descriptive comment for each query; do more than repeat the question.
- Complex statements should be broken over several lines as appropriate

Submitting your assignment

- Create a zip file containing your SQL script and report
- Submit assignment zip file via course website. Do not submit via email.

Evaluation Method

For each question you will receive the following marks:

- Execution: A variable number of marks will be assigned whether the command will run as provided.
- Accuracy: A variable number of marks will be assigned whether the question is answered correctly.

Additionally, the following marks will be assigned to the SQL script as a whole

- Structure: Up to 9 marks will be assigned for well-structured scripts that follow SQL standards. i.e. capitalized commands, lower case field names with underscores where needed, new lines, etc.
- Commenting: Up to 5 marks will be assigned for helpful, descriptive comments to be included throughout the script

If you have any questions, please do not hesitate to ask me.

Good luck.

Questions

1. Create a database for this assignment.
   Execution: 1.5 mark

2. Instruct MySQL Server to use the database you just created.
   Execution: 1.5 mark

3. In a separate SQL tab, open the *barrie_weather.sql* file provided. Under question 3, copy its contents into your assignment1 script, then run the commands to create the table (also called barrie_weather, and will be used to answer following questions) and then populate it with data using the provided insert statements.
   Execution: 2 mark

4. In the report only, explain how the provided CREATE TABLE statement could be improved. Please focus on the data types and parameters for this discussion.
   Written answer: 2 marks
   Bonus: Up to additional 3 marks

5. The provided INSERT statements in barrie_weather were generated automatically and load data very slowly because of the way that the script's commands have been structured. What is it about the structure of this script that causes the script to run slowly? In the report only, explain how you could restructure the commands if you were tasked with inserting the same data into barrie_weather.
   Written answer: 3 marks

6. Select all data for every hour of data where the temperature (temp) was greater than 26.5. Provide an additional comment in your SQL script that states the number of rows returned.
   Execution: 1.5 mark
   Accuracy: 2 marks

7. Select the date_time, temp, pressure_kPA and wind_spd_kmh for every hour of data where the temperature was 24 or less and the air pressure (pressure_kPa) was greater than 96.83. Provide an additional comment in your SQL script that states the number of rows returned.
   Execution: 1.5 mark
   Accuracy: 4 marks

8. Select the date and time (date_time), temperature (temp) and temperature flag (temp_flag) columns for every hour where temperature data has been marked as missing using an 'M' value in the temp_flag column.
   In the report, record the date and time.
   Hint: remember that single quotes are required for text values.
   Execution: 1.5 mark
   Accuracy: 2 marks

9. Return the same results as in Question 8, except construct the query to return both 'M' and 'Missing' values.
   Note: There are a number of ways of doing this, please use LIKE
   Execution: 1.5 mark
   Accuracy: 2 marks

10. Count (SELECT COUNT) the number of hours contained in this dataset, where the dew point temperature (dew_pt) is greater than one.
    Execution: 1.5 mark
    Accuracy: 2 mark

11. Count the number of hours that the wind was blowing (wind_spd_kmh is greater than 0) and from a direction ranging between north and east (wind_dir_10d is between 0 and 9)
    Execution: 1.5 mark
    Accuracy: 4 marks

12. In the report only, in Barrie what percent of the time are the conditions listed in Q11 true?
    Written answer: 1 mark

13. Select all data for every hour of data where either the temperature (temp) is below freezing, or wind chill (wind_chill) was below zero. Additionally, the wind speed flag (wind_spd_flag) must indicate that there were no problems with the wind speed instrument (i.e. no M or Missing values present)
BTW: an anemometer is an instrument that measures wind speed
    Execution: 1.5 mark
    Accuracy: 6 marks

14. Fix the M vs Missing problem in the temperature flag field by using UPDATE to change any instances of *Missing* values to just plain *M*.
    Execution: 1.5 mark
    Accuracy: 3 marks

15. Use an update statement to change all ' ' values in the temp_flag column to NULL (no quotes, just the letters N-U-L-L) this will change.
Hint: The values should end up looking like those in the humidex (hmdx) column.
FYI: The best way to store null data is using the NULL value. Look to the humidex (hmdx) column for an example of this being implemented properly.
    Execution: 1.5 mark
    Accuracy: 3 marks