

Hydroinformatics

Assignment 5

Using Structured Query Language

Due Date: October 27, 2016

Learning Objectives

1. Store, retrieve, and use data from important data models used in Hydrology such as ArcHydro, NetCDF, and the Observations Data Model (ODM)
2. Design and use relational databases to organize, store, and manipulate data
3. Query, aggregate, and pivot data using Structured Query Language (SQL), Excel, R, and other software systems

Computer and Data Requirements

1. You will need MySQL and MySQL Workbench installed to complete this assignment.
2. The data needed for this assignment have been provided in a zip file on the Canvas Website. You need to download the data and follow the instructions for setting up your Logan River ODM database before starting the assignment. Instructions are also available on the Canvas Website.

The Problem

You are working on a water quality study in the Logan River. One aspect of this study is to assess water temperature for potential impacts on cold-water fish species. The water quality standards for the State of Utah list the Logan River as Class 3A - protected for cold water species of game fish and other cold water aquatic life, including the necessary aquatic organisms in their food chain (see <http://www.rules.utah.gov/publicat/code/r317/r317-002.htm>). Given this, state water quality standards require that water temperatures in the river not exceed 20 degrees C to protect aquatic organisms. Your job is to perform some exploratory data analysis using the water temperature datasets in the Logan River ODM database and perform analyses that may identify potential water temperature impairment. In your analysis, you should write SQL queries on the database to assemble the following:

1. A table listing the period of record for water temperature measurements (e.g., begin and end date), the number of observations, and the overall minimum, maximum, and average values for each site at which quality controlled (QualityControlLevelID = 1) water temperature (VariableID = 57) data have been collected.
2. A table listing the total number of temperature observations, the number of observations greater than the water quality criterion value (i.e., 20 degrees C), and the overall percent exceedance of the water quality criterion value for each site at which quality controlled water temperature data have been collected.
3. A table for the Logan River at Mendon Road (SiteID = 2) listing the percent exceedance of the water quality standard for each month of the year.

4. A table listing the percent exceedance of the water quality standard for each site at which quality controlled data are available during the month of July, which is generally a critical period with low flows and elevated temperatures.

At which site(s) was the overall percent exceedance the greatest? Given that the state water quality standards are set to protect the health of aquatic organisms, at which site(s) is there most likely to be temperature related effects on fish populations? How does the percent exceedance change from month to month at the Mendon Road site? How does the percent exceedance in July compare to the overall percent exceedance?

Deliverable

Submit a one-page report that introduces the problem and addresses the questions above. Provide your tables as appendices to your one page write up. Additionally, provide an appendix that lists the code of the SQL queries that you used to generate your tables (indicate which query(ies) go with which tables). In your write up, please discuss what you learned.