Hydroinformatics

Assignment 5 Instructions

Prerequisites

To complete this exercise on your own computer, you must first install MySQL and MySQL Workbench. You should already have them installed, but if not, these software tools can be downloaded and installed using the instructions provided on the Canvas site with previous assignments.

Introduction

To complete Assignment 5, you need to create a new database in MySQL and load the data for the assignment. This new database will be an ODM database with the same structure as the one you used in Assignment 4. The difference is I'm giving you a database that has all of the quality controlled data values for the Logan River GAMUT network ready to load so we can practice writing SQL queries on a database that is populated with a good amount of data.

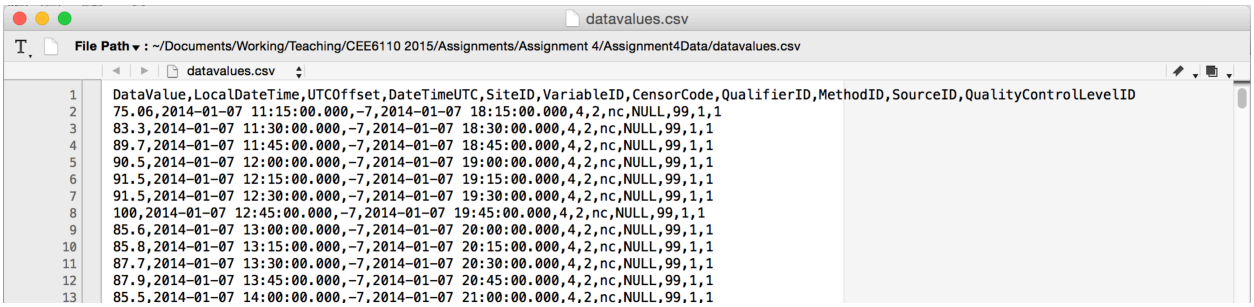
Creating the Logan River ODM Database

1. Download the zip file containing the data for this exercise from the Canvas website.
2. Extract the zip file to a convenient location on your computer. You should see a file called "createloganriverodm.sql" and a file called "datavalues.csv".
3. Open MySQL Workbench.
4. Click on the connection to your local MySQL Server instance to create a connection to your database server.
5. Click on the Create New Schema button on the toolbar at the top of the MySQL Workbench window. In the new schema tab, name your new schema "loganriverodm" and set the Default Collation to "utf8 – utf8_unicode_ci". Then click the "Apply" button.
6. In the Review SQL Script form that pops up, click the "Apply" button. Click the "Close" button when you receive a message saying that the script was successfully applied to the database.
7. On the MySQL Workbench "File" drop down menu select "Open SQL Script". In the file browser that pops up, select the file "createloganriverodm.sql" from the location where you unzipped the data for this exercise. The script will now be open in a new script window in MySQL Workbench.
8. Take a minute to look over the script that I have prepared. It is very similar to the one we ran for the last assignment to create a blank ODM database. The biggest difference is that toward the end of the script it actually inserts data into most of the tables in ODM.

NOTE: I created this script by dumping data out of the production ODM database for the Logan River GAMUT sites. A SQL script is one way to transfer a relational database from one system to another system.

9. Click the lightning bolt button on the script editor tab to execute the script. You will see messages pop up in your Action Output pane as the script creates the tables and adds data to them. When the script has completed, you will see a message at the very bottom of the MySQL Workbench window stating that the query is complete.

10. You may have to right click on the "loganriverodm" database and click "Referesh All" to get the tables to show up. At this point, you should be able to right click on most of the tables and choose "Select Rows" to see the data that was loaded to all of the tables.
11. Now you need to load the data values I have provided as a CSV file into the DataValues table. Use the "LOAD DATA LOCAL INFILE" functionality that we used in the last assignment in a SQL script to load the data from the CSV file. The file is quite large (~ 150 MB) and contains about 2 million DataValues so it might take a couple of minutes to load (it took about 40 seconds on my machine). Don't forget to set the path to your "datavalues.csv" file correctly. You are loading data into the DataValues table in the ODM database, and you need to list the Attributes (the columns) from the file that you are loading. Here's a snippet of the header of the file in case you have trouble opening it to see what columns it contains:



12. As a last step, you should execute the stored procedure in your database to update the SeriesCatalog table. Open a new script window by clicking on the new script button on the toolbar. In the new script window, type the following code:

```
USE loganriverodm;
CALL sp_UpdateSeriesCatalog();
```

13. Execute the script by clicking the lightning bolt icon.
14. Once the script is complete, you should be able to right click on the SeriesCatalog table and choose "Select Rows."
15. You will now see a tabular view of the series catalog table with a record for each of the time series that you loaded from the file I gave you.

NOTE: Once you have completed these steps your database is set up and ready to go for Assignment 5. Feel free to use ODM Tools Python to visualize and explore the data you just loaded to the database.