

CSE 111 – DATABASE SYSTEMS

Lab 2

In this lab session you are required to bulk-load TPC-H data from CSV files into the SQLite database you created in Lab 1. This step is required in order to be able to run meaningful queries on the database.

In order to complete the lab you have to perform the following tasks:

1. Log in to your GitLab account.
2. Explore the folders and files in the Lab 2 repo.
3. Create a merge request for the **Instructions** issue. This is done from the **Issues** tab. The result of the merge request is a new branch that copies the files from **master**.
4. Clone the repo to your local machine or the remote lab machine. You can choose to directly clone the branch for the merge request, or the **master** and then checkout the merge request branch.
5. There is a CSV file corresponding to each TPC-H table in the **data/** folder. Inspect the content of these files to identify the separator between attributes/columns.
6. Write a **SQL** bulk-loading statement for every table as shown in the lecture on database modification operations. All the **SQL** statements have to be written in the file **load-tpch.sql**. This is the only file you have to edit in this lab.
7. Once the code in **load-tpch.sql** is executed, the tables in the database have to contain the same data as in the files.
8. You can check the correctness of your loading code by executing the command **make run** in the terminal. You have to be in the main lab folder. The expected output is available in **results/*.res**. The output produced by your code is available in **output/*.out**. They have to match for every query, e.g., **1.res** has to match with **1.out**.
9. Commit the changes to **load-tpch.sql** and then push to the GitLab server.
10. Check the output of the pipeline under the **CI / CD** tab to see if your push has passed all the tests.
11. In case there are any errors, repeat the process from step 6.

The score for the lab is assigned based on passing the test cases and the commit/push history. The instructor and the TAs have access to the GitLab repos.