CSE 111 – DATABASE SYSTEMS
Lab 2: Load Data Into TPCH Database

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. 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.

2. 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.

3. Once the code in `load-tpch.sql` is executed, the tables in the database have to contain the same data as in the files.

4. You can check the correctness of your loading code by executing the command `./test.sh` 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`.

5. In case there are any errors, repeat the process from step 2.

You have to submit the modified file `load-tpch.sql` in CatCourses. This is the only file you have to upload.

The score for the lab is assigned based on how many test cases are passed.