

## CSE 111 – DATABASE SYSTEMS

### Lab 6

In this lab session you have to write 15 SQL queries for the TPC-H database created and populated in the previous labs. The queries are the following (1 point per query):

1. Find the total quantity (`l_quantity`) of line items shipped per month (`l_shipdate`) in 1995. Hint: check function `strftime` to extract the month/year from a date.
2. Find the number of customers who had at least three orders in November 1995 (`o_orderdate`).
3. Find how many parts are supplied by exactly two suppliers from UNITED STATES.
4. Find how many suppliers from UNITED STATES supply more than 40 different parts.
5. Find how many suppliers supply the most expensive part (`p_retailprice`).
6. Find the supplier-customer pair(s) with the least expensive (`o_totalprice`) order(s) completed (F in `o_orderstatus`). Print the supplier name, the customer name, and the total price.
7. Find how many suppliers have less than 50 distinct orders from customers in GERMANY and FRANCE together.
8. Find how many distinct customers have at least one order supplied exclusively by suppliers from AMERICA.
9. Find the distinct parts (`p_name`) ordered by customers from AMERICA that are supplied by exactly 3 suppliers from ASIA.
10. Find the region where customers spend the smallest amount of money (`l_extendedprice`) buying items from suppliers in the same region.
11. Find the nation(s) with the smallest number of customers.
12. Find the nation(s) having customers that spend the smallest amount of money (`o_totalprice`).
13. Find the nation(s) with the most developed industry, i.e., selling items totaling the largest amount of money (`l_extendedprice`) in 1994 (`l_shipdate`).
14. Compute, for every country, the value of economic exchange, i.e., the difference between the number of items from suppliers in that country sold to customers in other countries and the number of items bought by local customers from foreign suppliers in 1994 (`l_shipdate`).
15. Compute the change in the economic exchange for every country between 1994 and 1996. There should be two columns in the output for every country: 1995 and 1996. Hint: use `CASE` to select the values in the result.

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

1. Write the SQL statement corresponding to every query in the file `test/x.sql`, where `x` is the number of the query above. Every query goes into its separate file. These are the only files you have to modify and submit in this assignment.
2. You can check the correctness of your queries 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/x.res`, where `x` is the number of the query. The output produced by your code is available in `output/x.out`. They have to match for every query, e.g., `1.res` has to match with `1.out`.

3. In case there are any errors, repeat the process from step 1 for the incorrect queries.
4. The submission consists of a compressed **zip** file that contains the files in the **test** folder. The name of the file has to be **lab-6.zip**. When you create the file, include the folder **test** into the compression, not every file **test/x.sql** separately.