

CSE 111 – DATABASE SYSTEMS

Lab 5

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

1. How many customers are not from EUROPE or AFRICA?
2. How many suppliers in every region have more balance in their account than the average account balance of their own region?
3. For the line items ordered in May 1995 (o_orderdate), find the largest discount that is less than the average discount among all the orders.
4. How many customers and suppliers are in every country from EUROPE?
5. For parts whom type contains STEEL, return the name of the supplier from AMERICA that can supply them at minimum cost (ps_supplycost), for every part size. Print the supplier name together with the part size and the minimum cost.
6. Based on the available quantity of items, who is the manufacturer p_mfgr of the most popular item (the more popular an item is, the less available it is in ps_availqty) from Supplier#000000053?
7. For every order priority, count the number of parts ordered in 1996 and received earlier (l_receiptdate) than the commit date (l_commitdate). List the results in descending priority order.
8. Count the number of distinct suppliers that supply parts whom type contains MEDIUM POLISHED and have size equal to any of 3, 23, 26, and 49.
9. Count the number of supplied parts that have total value (ps_supplycost*ps_availqty) in the top 3% values across all the supplied parts and are supplied by suppliers from CANADA. Hint: Use the LIMIT keyword.
10. How many customers from every region have never placed an order and have more than the average account balance?
11. Find the highest value items (l_extendedprice*(1-l_discount)) not shipped as of October 2, 1994. Print the name of the part corresponding to these items.
12. What is the total supply cost for parts less expensive than \$1000 (p_retailprice) shipped in 1996 (l_shipdate) by suppliers who did not supply any item with an extended price less than 2000 in 1995?
13. Count the number of orders made in the fourth quarter of 1996 in which at least one item was received by a customer later than its commit date. List the count of such orders for every order priority sorted in ascending priority order.
14. For any two regions, find the gross discounted revenue (l_extendedprice*(1-l_discount)) derived from line items in which parts are shipped from a supplier in the first region to a customer in the second region in 1995 and 1996. List the supplier region, the customer region, the year (l_shipdate), and the revenue from shipments that took place in that year. Order the answers by supplier region, customer region, and year.
15. The market share for a given nation within a given region is defined as the fraction of the revenue from the line items ordered by customers in the given region that are supplied by suppliers from the given nation. The revenue of a line item is defined as l_extendedprice*(1-l_discount). Determine the market share of UNITED STATES in EUROPE in 1996 (l_shipdate).

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 5 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. Write the **SQL** statement corresponding to each query in the file **test/x.sql**, where **x** is the number of the query above. Each query goes into its separate file. These are the only files you have to modify and commit in this assignment.
6. You can check the correctness of your queries 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/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**.
7. Commit the changes to the query files and then push to the GitLab server.
8. Check the output of the pipeline under the **CI / CD** tab to see if your push has passed all the tests.
9. In case there are any errors, repeat the process from step 5.

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.