

**CSCI235 – Database Systems
2024 S2
Implementation Task 2
Due on 12 May 2024 by 9:00 pm SGT**

Scope

The Implementation of Task 2 is related to the contents of topic on indexing.

This Implementation is due by Sunday, 12 May 2024, 9:00 pm Singapore time.
This task is worth 3% of the total assessment for the subject.

Only electronic submission through Moodle at:

<https://moodle.uowplatform.edu.au/>

is accepted. All email submission will be deleted and mark 0 ("zero") will be awarded.

For all the implemented tasks, your report or output must include a listing of all PL/SQL statements processed. To achieve that put the following SQL*Plus commands in all your scripts:

```
SPOOL file-name  
SET ECHO ON  
SET FEEDBACK ON  
SET LINESIZE 100  
SET PAGESIZE 200  
SET SERVEROUTPUT ON
```

at the beginning of SQL script and

```
SPOOL OFF
```

at the end of SQL script.

The submission procedure is explained at the end of this specification.

Specification

Task 1 (1.5 marks)

Indexing

Using the relational table LINEITEM of the sample database TPCHR, for each one of the queries listed below:

- i. Find all the discount (l_discount) of all the items that are shipped (l_shipdate) most recently. Hint. Most recently mean the latest shipment date.
 - ii. Find the total number of items shipped by air (l_shipmode) in 1998 (l_shipdate).
 - iii. Find the order number (l_orderkey) and item number (l_linenum) that have the highest discount (l_discount).
 - iv. Find the total number of item per line status (l_linestatus). List the line status and the total items per line status.
 - v. Find the order key (l_orderkey), line item number (l_linenum), line status (l_linestatus), shipment date (l_shipdate) and shipment mode (l_shipmode) of all orders with the order number (l_orderkey) 1795718, 1799046, and 1794626.
- a) Construct an SQL statement that produces the required output specified in the statement. **(0.5 marks)**
- b)** Find the best possible indexing of a relational table LINEITEM. The best possible indexing means that a database system will compute a query with an index proposed by you using the smallest number of read block operations. Note that you can create only one index per query, and there is no need to find indexes that speed up the processing of more than one query. Use the explain plan and show plan statements to justify your solutions. **(0.5 marks)**

Deliverables

A file **It2Solution1.sql** with CREATE INDEX statements that improve the performance of the queries listed (i, ii, iii, iv, and v above) and the execution plan generated, **It2Solution1Output.lst**.

Please remember that you must consider each one of the queries as an individual case! Please remember that all relational tables are large enough to make full table scans more time consuming than accessing the tables through an index! It means that any solution in which an index is not used for query processing is incorrect.

Task 2 (1.5 marks)

Indexing

Using the relational table ORDERS of the sample database TPCHR and the following index created, construct an appropriate SQL statement and satisfies the following requirement:

create index A1Task3Idx on orders(o_orderdate, o_clerk, o_totalprice);

Find SELECT statements that will use the index in the following ways:

- a) Execution of the first SELECT statement must traverse the index vertically and it must not access a relational table ORDERS. **(0.3 mark)**
- b) Execution of the second SELECT statement must first traverse the index vertically followed by horizontal scan at the leaf level. The retrieval operation must not access a relational table ORDERS. **(0.3 mark)**
- c) Execution of the third SELECT statement must traverse the leaf level of the index horizontally and it must not access the relational table ORDERS. **(0.3 mark)**
- d) Execution of the fourth SELECT statement must traverse the index vertically and it must access a relational table ORDERS. **(0.3 mark)**
- e) Execution of the fifth SELECT statement must first traverse the index vertically followed by horizontal scan at the leaf level. The retrieval operation must access a relational table ORDERS. **(0.3 mark)**

Deliverables

A file **It2Solution2Output.lst** and **It2Solution2.sql** with the SELECT statements for each one of the five cases described above.

Submissions

This implementation task is due by 9:00 pm (21:00 hours) Sunday, 12 May 2024, Singapore time.

Submit the files **It2Solution1.sql**, **It2Solution1Output.lst**, **It2Solution2.sql**, **It2Solution2Output.lst** through Moodle in the following way. You can also zip them into a zipped file. You can name the zipped file as It2Solutions.zip:

- 1) Access Moodle at **<http://moodle.uowplatform.edu.au/>**
- 2) To login use a Login link located in the right upper corner the Web page or in the middle of the bottom of the Web page
- 3) When successfully logged in, select a site CSCI235 (SP224) Database Systems
- 4) Scroll down to the section "Implementation Task 2 Submission link" and click at the link.
- 5) Click at a button Add Submission
- 6) Move the file that you want to submit into an area of submission. If you have more than one files to submit, you can drag and drop the remaining files here to add them. You can also use a link *Add...*
- 7) Click at the button "Save changes",
- 8) Click at the check box to confirm authorship of your submission and click at the button "Continue".
- 9) When you are satisfied, remember to click at a button Submit assignment.

A policy regarding late submissions is included in the subject outline. Only one submission per student is accepted.

Implementation Task 2 is an individual assignment, and it is expected that all its tasks will be solved individually without any cooperation with the other students. Plagiarism is treated seriously. Students involved will likely receive zero. If you have any doubts, questions, etc. please consult your lecturer or tutor during lab classes or over e-mail.