**Project: DBMS Implementation – Part III (50 points)**

**Due Dates: Dec 3, 2021**

Description:

There are 2 more SQL features you need to implement on top of your project spec. There will be a separate checklist given for these 2 features.

Specification:

Statement syntax and its function:

Note: { ... } means the content of ... within {} pair can be repeated separated by a comma ','.

Note: [ ... ] means the content within [] is optional.

Note: The '{', '}', '[', ']' symbols are NOT part of the syntax.

1) GROUP BY

SELECT [column\_name,] <aggregate>(column\_name) FROM table\_name

[ WHERE column\_name <condition> [(AND | OR) column\_name <condition>] ]

[ GROUP BY column\_name ]

[ ORDER BY column\_name [DESC] ]

- <condition> : <relational\_operator> data\_value

- <condition> : IS NULL

- <condition> : IS NOT NULL

- Return 0 or more rows of data which matches the search condition.

- <aggregate> can be SUM, AVG, COUNT

- SUM & AVG are only valid on integer column. The \* symbol is not a valid substitution.

- COUNT can be used in any column type or \*, it always count the # of rows depending on the condition.

- The GROUP BY column\_name must be the same as the [column\_name] specified in the select list.

2) NATURAL JOIN

SELECT { [table\_name.]column\_name } FROM table\_name [NATURAL JOIN table\_name]

[ WHERE [table\_name.]column\_name <condition> [(AND | OR) [table\_name.]column\_name <condition>] ]

[ ORDER BY column\_name [DESC] ]

- <condition> : = column\_name

- <condition> : <relational\_operator> data\_value

- <condition> : IS NULL

- <condition> : IS NOT NULL

- Return 0 or more rows of data which matches the search condition.

- The column\_name in the select list can be replaced by the \* symbol meaning all the columns.

- If the ORDER BY clause is not there, then display the records in the storage order.

- <aggregate> can be SUM, AVG, COUNT

- SUM & AVG are only valid on integer column. The \* symbol is not a valid substitution.

- COUNT can be used in any column type or \*, it always count the # of rows depending on the condition.

- Implement the basic NLJ algorithm for INNER JOIN.

- Need to support columns in select list from different tables by adding the table\_name qualifier.

CHECKLIST:

1. \_\_\_\_\_ Different test scenarios for GROUP BY:
   1. Group By with and without a WHERE clause
   2. Group By with and without an ORDER BY clause.
   3. Group By with different combination of SUM, AVG, COUNT with and without WHERE clause
2. \_\_\_\_\_ Different test scenarios for NATURAL JOIN
   1. Join 2 tables with 1, 2, and 3 shared columns names
   2. Join 2 tables with a \* in the SELECT list.
   3. Join 2 tables with SELECT list columns from 2 different tables using the table\_name qualifier

- There is a 15 points extra credit if you extend the NATURAL JOIN feature to 3 tables. (e.g. table1 NATURAL JOIN table2 NATURAL JOIN table3

- In order to get extra credit, you must finished both features first and submit before the “Due” date.