

# P3: Database Operations

---

Group Members: Kayton Fletcher

The bulk of the implementation for project 3 is within the `relational_ops` folder within the project directory. This folder is where I made a class (and therefore an associated header and `.cc` file) for every operation required to perform database queries, as requested in the specification.

For this project I overhauled `test.h` to remove many of the bad practices and annoyances it had. To do this I used `inline`, so this project *must be compiled with c++17*. Similarly, changes were made to project 3's `test.cc` to work with the changed header file.

I also greatly restructured the file layout into directories to organize the unwieldy number of files in the project. I did this before reading the mention in the announcement to not do this. This makefile should ensure this still works with a script, as it compiles the `test.out` executable.

Each operation created spawns a worker thread that can be joined by calling the specific `Relation::WaitUntilDone()` function. Each operation also has an associated `Run` function that spawns the thread and performs a calculation. The operations implemented include `Join`, `SelectPipe`, `SelectFile`, `DuplicateRemoval`, `GroupBy`, `Project`, and `Sum`. There is also an operation to write records out to a file in a nice format, namely `WriteOut`.

## Assumptions

It is assumed that both `bison` and `flex` are installed to compile the parser needed. Additionally, `gtest` must be installed as a library that can be linked on the system, as I only specify `-lgtest` within the `Makefile` and will not be providing the program with my submission.

The needed files `catalog`, `Lexer.l`, `LexerFunc.l`, `ParserFunc.y` and `Parser.y` have been included in the zip file. They are found within the `dependencies` subdirectory.

The variables `dbfile_dir`, `tpch_dir`, and `catalog_path` now have their values derived from the `test.cat` file found in the projects root directory. If `test.h` is not present, default values are given for the static variables in `test.h`.

Finally, many adjustments were made to the starter code provided, and as such I am expecting the files I have provided to be used when compiling my DB classes.

## Compiling The Programs

To compile `test.out`, the program that runs `test.cc` as provided, run the following command

```
make test.out
```

To compile `gtest.out`, the program that runs the unit tests I wrote, run the command

```
make gtests.out
```

To compile `a2test.out`, the program that enables the creation of heap and sorted database \*.bin files

```
make a2test.out
```

No `main.cc` was submitted with this project, other than the one in the `gtests` folder.

## Running The Programs

Running the programs is very straight forward. Treat each as an executable.

`./test.out` and `./gtests.out`.

Neither executable requires user input. However, `test.out` must have a command line argument for which test you wish to run. `gtests.out` does need a valid location for the HeapDB for the `nation` and `region` relations.

The unit tests created with gtest are all located in `gtests/main.cc`. The gtests created test the relational operations created.

To run the program that enables database creation, simply run `./a2test.out`. From there you may choose option 1 to create a sorted database or option 2 to create a heap database. For either, you must specify the relation you wish to create. Both use the locations specified in `test.cat` to retrieve and write files. For sorted files, you must also specify the run length and sorting order desired.

## `./runTestCases.sh` Output

```
51 int: [9022], string: [peru misty sandy dark drab], double: [931.02]
52 int: [10021], string: [blush steel green sienna snow], double: [931.02]
53 int: [11020], string: [plum khaki powder beige peru], double: [931.02]
54
55
56 query2 returned 12 records
57
58
59 *****
60
61 ** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
62 catalog location: catalog
63 tpch files dir: ./p1/tables/
64 heap files dir: ./heap_dbs/
65
66
67 double: [9.24623e+07]
68
69
70 query3 returned 1 records
71
72
73 *****
74
75 ** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
76 catalog location: catalog
77 tpch files dir: ./p1/tables/
78 heap files dir: ./heap_dbs/
79
80
81 query4
82 double: [4.00421e+08]
83 query4 returned 1 recs
84
85
86 *****
87
88 ** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
89 catalog location: catalog
90 tpch files dir: ./p1/tables/
91 heap files dir: ./heap_dbs/
92
93
94 query5 finished..output written to file ps.w.tmp
95
```

## Gtest Output

```

    use never according to the quickly regular pint]
int: [3], string: [EUROPE], string: [ly final courts cajole furiously final excuse
], int: [23], string: [UNITED KINGDOM], int: [3], string: [eans boost carefully sp
ecial requests. accounts are. carefull]
int: [4], string: [MIDDLE EAST], string: [uickly special accounts cajole carefully
blithely close requests. carefully final asymptotes haggle furiousl], int: [13],
string: [JORDAN], int: [4], string: [ic deposits are blithely about the carefully
regular pa]
int: [4], string: [MIDDLE EAST], string: [uickly special accounts cajole carefully
blithely close requests. carefully final asymptotes haggle furiousl], int: [4], s
tring: [EGYPT], int: [4], string: [y above the carefully unusual theodolites. fina
l dugouts are quickly across the furiously regular d]
int: [4], string: [MIDDLE EAST], string: [uickly special accounts cajole carefully
blithely close requests. carefully final asymptotes haggle furiousl], int: [11],
string: [IRAQ], int: [4], string: [nic deposits boost atop the quickly final reques
ts? quickly regula]
int: [4], string: [MIDDLE EAST], string: [uickly special accounts cajole carefully
blithely close requests. carefully final asymptotes haggle furiousl], int: [20],
string: [SAUDI ARABIA], int: [4], string: [ts. silent requests haggle. closely exp
ress packages sleep across the blithely]
int: [4], string: [MIDDLE EAST], string: [uickly special accounts cajole carefully
blithely close requests. carefully final asymptotes haggle furiousl], int: [10],
string: [IRAN], int: [4], string: [efully alongside of the slyly final dependencie
s. ]
[      OK ] Join.SimpleJoin (4 ms)
[ RUN      ] Join.JoinWithSum

** IMPORTANT: MAKE SURE THE INFORMATION BELOW IS CORRECT **
catalog location:      catalog
tpch files dir:        ./p1/tables/
heap files dir:        ./heap_dbs/

int: [300]
[      OK ] Join.JoinWithSum (5 ms)
[-----] 2 tests from Join (9 ms total)

[-----] Global test environment tear-down
[=====] 5 tests from 3 test suites ran. (24 ms total)
[ PASSED ] 5 tests.

```