# Distribution of Parameters for Generated SQL Queries on the TPC-H Database

```
## Scenario: 1-6jp_rjjp__g__h0-3ss_h0-3nss

## Scale factor: 1 GB

## The 1000-query pack was generated on: 2023-11-12

## First column (‘Query parameter or value‘) displays either:

##  - a parameter name, or

##  - a value of the current parameter (only for low-cardinality parameters).

## Second column (‘Descriptive statistics‘) displays:

##  - for low cardinality parameters: frequency and proportion from total (of 1000 queries)

##  - for higher cardinality parameters: [min, median, max]  mean / SD
```

| **Query parameter or value** | **Descriptive statistics** |
|---|---|
| SELECT_n_of_columns | [1, 8, 174] 15 / 23 |
| SELECT_n_of_non_aggr_func___FLOOR | |
| *0* | 931 (93%) |
| *1* | 62 (6.2%) |
| *2* | 4 (0.4%) |
| *3* | 1 (0.1%) |
| *4* | 1 (0.1%) |
| *5* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LTRIM | |
| *0* | 637 (64%) |
| *1* | 245 (25%) |
| *2* | 79 (7.9%) |
| *3* | 19 (1.9%) |
| *4* | 13 (1.3%) |
| *5* | 4 (0.4%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *6* | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___MONTH | |
| *0* | 806 (81%) |
| *1* | 133 (13%) |
| *2* | 33 (3.3%) |
| *3* | 18 (1.8%) |
| *4* | 3 (0.3%) |
| *5* | 4 (0.4%) |
| *7* | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___RTRIM | |
| *0* | 633 (63%) |
| *1* | 261 (26%) |
| *2* | 61 (6.1%) |
| *3* | 21 (2.1%) |
| *4* | 11 (1.1%) |
| *5* | 7 (0.7%) |
| *6* | 2 (0.2%) |
| *7* | 3 (0.3%) |
| *8* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___SQRT | |
| *0* | 800 (80%) |
| *1* | 147 (15%) |
| *2* | 27 (2.7%) |
| *3* | 13 (1.3%) |
| *4* | 4 (0.4%) |
| *5* | 5 (0.5%) |
| *6* | 1 (0.1%) |
| *7* | 2 (0.2%) |
| *8* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___UPPER | |
| *0* | 647 (65%) |
| *1* | 242 (24%) |
| *2* | 71 (7.1%) |
| *3* | 21 (2.1%) |
| *4* | 13 (1.3%) |
| *5* | 5 (0.5%) |
| *7* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___YEAR | |
| *0* | 822 (82%) |
| *1* | 128 (13%) |
| *2* | 33 (3.3%) |
| *3* | 11 (1.1%) |
| *4* | 4 (0.4%) |
| *5* | 1 (0.1%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___SUBSTR | |

| **Query parameter or value** | **Descriptive statistics** |
|---|---|
| *0* | 621 (62%) |
| *1* | 252 (25%) |
| *10* | 1 (0.1%) |
| *2* | 85 (8.5%) |
| *3* | 26 (2.6%) |
| *4* | 10 (1.0%) |
| *5* | 3 (0.3%) |
| *6* | 1 (0.1%) |
| *8* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___DAY | |
| *0* | 891 (89%) |
| *1* | 82 (8.2%) |
| *2* | 21 (2.1%) |
| *3* | 6 (0.6%) |
| SELECT_n_of_non_aggr_func___ROUND | |
| *0* | 911 (91%) |
| *1* | 68 (6.8%) |
| *2* | 17 (1.7%) |
| *3* | 4 (0.4%) |
| SELECT_n_of_non_aggr_func___TRUNC | |
| *0* | 927 (93%) |
| *1* | 66 (6.6%) |
| *2* | 4 (0.4%) |
| *3* | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___ABS | |
| *0* | 910 (91%) |
| *1* | 79 (7.9%) |
| *2* | 7 (0.7%) |
| *3* | 3 (0.3%) |
| *4* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___DOW | |
| *0* | 856 (86%) |
| *1* | 107 (11%) |
| *2* | 22 (2.2%) |
| *3* | 7 (0.7%) |
| *4* | 6 (0.6%) |
| *5* | 1 (0.1%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOWER | |
| *0* | 634 (63%) |
| *1* | 260 (26%) |
| *2* | 68 (6.8%) |
| *3* | 22 (2.2%) |
| *4* | 7 (0.7%) |
| *5* | 5 (0.5%) |
| *6* | 1 (0.1%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *7* | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___LOG | |
| *0* | 848 (85%) |
| *1* | 118 (12%) |
| *2* | 22 (2.2%) |
| *3* | 7 (0.7%) |
| *4* | 3 (0.3%) |
| *5* | 1 (0.1%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_all_non_aggr_func | [0.0, 3.0, 52.0] 4.6 / 6.7 |
| SELECT_n_of_aggr_func___COUNT | [0.00, 0.00, 14.00] 0.82 / 1.94 |
| SELECT_n_of_aggr_func___COUNT_DISTINCT | [0.00, 0.00, 17.00] 0.88 / 2.01 |
| SELECT_n_of_aggr_func___MIN | [0.00, 0.00, 11.00] 0.79 / 1.87 |
| SELECT_n_of_aggr_func___SUM | |
| *0* | 924 (92%) |
| *1* | 55 (5.5%) |
| *2* | 15 (1.5%) |
| *3* | 3 (0.3%) |
| *4* | 3 (0.3%) |
| SELECT_n_of_aggr_func___MAX | [0.00, 0.00, 12.00] 1.01 / 2.14 |
| SELECT_n_of_aggr_func___AVG | |
| *0* | 930 (93%) |
| *1* | 41 (4.1%) |
| *2* | 20 (2.0%) |
| *3* | 8 (0.8%) |
| *4* | 1 (0.1%) |
| SELECT_n_of_all_aggr_func | [0.00, 3.00, 17.00] 3.71 / 3.19 |
| FROM_n_of_join_paths | |
| *1* | 153 (15%) |
| *2* | 157 (16%) |
| *3* | 168 (17%) |
| *4* | 173 (17%) |
| *5* | 174 (17%) |
| *6* | 175 (18%) |
| FROM_n_of_super_joins___LEFT | |
| *0* | 532 (53%) |
| *1* | 329 (33%) |
| *2* | 113 (11%) |
| *3* | 24 (2.4%) |
| *4* | 2 (0.2%) |
| FROM_n_of_super_joins___RIGHT | |
| *0* | 547 (55%) |
| *1* | 310 (31%) |
| *2* | 121 (12%) |
| *3* | 17 (1.7%) |
| *4* | 5 (0.5%) |

| **Query parameter or value** | **Descriptive statistics** |
|---|---|
| FROM_n_of_super_joins___FULL | |
| *0* | 533 (53%) |
| *1* | 315 (32%) |
| *2* | 115 (12%) |
| *3* | 33 (3.3%) |
| *4* | 3 (0.3%) |
| *5* | 1 (0.1%) |
| FROM_n_of_joins___INNER | [0.00, 3.00, 13.00] 3.45 / 2.51 |
| FROM_n_of_joins___RIGHT | [0.00, 3.00, 15.00] 3.42 / 2.50 |
| FROM_n_of_processed_rows | [5, 9,300,646, 43,923,626] 12,229,603 / 10,138,395 |
| WHERE_n_of_predicates | [0.00, 4.00, 14.00] 4.00 / 2.94 |
| WHERE_n_of_attribs_of_type___character | |
| *0* | 628 (63%) |
| *1* | 246 (25%) |
| *2* | 96 (9.6%) |
| *3* | 20 (2.0%) |
| *4* | 8 (0.8%) |
| *5* | 2 (0.2%) |
| WHERE_n_of_attribs_of_type___character_varying | |
| *0* | 316 (32%) |
| *1* | 314 (31%) |
| *2* | 186 (19%) |
| *3* | 124 (12%) |
| *4* | 41 (4.1%) |
| *5* | 15 (1.5%) |
| *6* | 4 (0.4%) |
| WHERE_n_of_attribs_of_type___integer | |
| *0* | 372 (37%) |
| *1* | 274 (27%) |
| *2* | 202 (20%) |
| *3* | 100 (10%) |
| *4* | 35 (3.5%) |
| *5* | 10 (1.0%) |
| *6* | 6 (0.6%) |
| *7* | 1 (0.1%) |
| WHERE_n_of_attribs_of_type___numeric | |
| *0* | 575 (58%) |
| *1* | 260 (26%) |
| *2* | 125 (13%) |
| *3* | 28 (2.8%) |
| *4* | 9 (0.9%) |
| *5* | 3 (0.3%) |
| WHERE_n_of_attribs_of_type___date | |
| *0* | 771 (77%) |
| *1* | 189 (19%) |
| *2* | 31 (3.1%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *3* | 8 (0.8%) |
| *6* | 1 (0.1%) |
| WHERE_n_of_pkey_attribs | |
| *0* | 551 (55%) |
| *1* | 285 (29%) |
| *2* | 114 (11%) |
| *3* | 37 (3.7%) |
| *4* | 13 (1.3%) |
| WHERE_n_of_connect_OR | [0.00, 2.00, 12.00] 2.34 / 2.44 |
| WHERE_n_of_operators___between | |
| *0* | 500 (50%) |
| *1* | 305 (31%) |
| *2* | 145 (15%) |
| *3* | 39 (3.9%) |
| *4* | 10 (1.0%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_operators___greater_or_less | |
| *0* | 319 (32%) |
| *1* | 271 (27%) |
| *2* | 193 (19%) |
| *3* | 110 (11%) |
| *4* | 67 (6.7%) |
| *5* | 29 (2.9%) |
| *6* | 7 (0.7%) |
| *7* | 3 (0.3%) |
| *8* | 1 (0.1%) |
| WHERE_n_of_operators___in | |
| *0* | 513 (51%) |
| *1* | 302 (30%) |
| *2* | 125 (13%) |
| *3* | 43 (4.3%) |
| *4* | 13 (1.3%) |
| *5* | 4 (0.4%) |
| WHERE_n_of_operators___like | |
| *0* | 763 (76%) |
| *1* | 210 (21%) |
| *2* | 27 (2.7%) |
| WHERE_n_of_non_aggr_func___LOG | |
| *0* | 921 (92%) |
| *1* | 74 (7.4%) |
| *2* | 5 (0.5%) |
| WHERE_n_of_non_aggr_func___ABS | |
| *0* | 952 (95%) |
| *1* | 47 (4.7%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___DOW | |

| **Query parameter or value** | **Descriptive statistics** |
|---|---|
| *0* | 961 (96%) |
| *1* | 37 (3.7%) |
| *2* | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___SQRT | |
| *0* | 890 (89%) |
| *1* | 100 (10%) |
| *2* | 9 (0.9%) |
| *3* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___MONTH | |
| *0* | 956 (96%) |
| *1* | 42 (4.2%) |
| *2* | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___FLOOR | |
| *0* | 955 (96%) |
| *1* | 44 (4.4%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___DAY | |
| *0* | 956 (96%) |
| *1* | 42 (4.2%) |
| *2* | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___ROUND | |
| *0* | 958 (96%) |
| *1* | 42 (4.2%) |
| WHERE_n_of_non_aggr_func___TRUNC | |
| *0* | 961 (96%) |
| *1* | 38 (3.8%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___YEAR | |
| *0* | 961 (96%) |
| *1* | 39 (3.9%) |
| WHERE_n_of_all_non_aggr_func | |
| *0* | 618 (62%) |
| *1* | 250 (25%) |
| *2* | 99 (9.9%) |
| *3* | 26 (2.6%) |
| *4* | 7 (0.7%) |
| GROUP_BY_n_of_columns | [0.00, 2.00, 12.00] 3.05 / 2.85 |
| HAVING_n_of_main_predicates | [0.00, 4.00, 9.00] 3.49 / 2.58 |
| HAVING_n_of_main_predicates___scalar_subquery | |
| *0* | 429 (43%) |
| *1* | 167 (17%) |
| *2* | 201 (20%) |
| *3* | 203 (20%) |
| HAVING_n_of_main_predicates___non_scalar_subquery | |
| *0* | 441 (44%) |
| *1* | 168 (17%) |

| **Query parameter or value** | **Descriptive statistics** |
|---|---|
| *2* | 190 (19%) |
| *3* | 201 (20%) |
| HAVING_n_of_subqueries___scalar_subquery | [0.0, 1.0, 19.0] 2.5 / 3.6 |
| HAVING_n_of_subqueries___non_scalar_subquery | |
| *0* | 441 (44%) |
| *1* | 168 (17%) |
| *2* | 190 (19%) |
| *3* | 201 (20%) |
| HAVING_n_of_processed_rows_by_subqueries | [0, 13,880,260, 647,387,190] 45,482,539 / 75,614,053 |
| ORDER_BY_n_of_columns | |
| *0* | 302 (30%) |
| *1* | 355 (36%) |
| *2* | 214 (21%) |
| *3* | 84 (8.4%) |
| *4* | 20 (2.0%) |
| *5* | 14 (1.4%) |
| *6* | 9 (0.9%) |
| *7* | 2 (0.2%) |
| limit | [1, 498, 1,000] 505 / 283 |
| offset | [0, 0, 999] 226 / 309 |