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

## Scenario: 1-6jp_rjjp__g__h0ss_h0nss

## Scale factor: 2 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, 9, 178] 16 / 24 |
| SELECT_n_of_non_aggr_func___FLOOR | |
| 0 | 908 (91%) |
| 1 | 81 (8.1%) |
| 2 | 11 (1.1%) |
| SELECT_n_of_non_aggr_func___LOG | |
| 0 | 847 (85%) |
| 1 | 112 (11%) |
| 2 | 21 (2.1%) |
| 3 | 10 (1.0%) |
| 4 | 4 (0.4%) |
| 5 | 3 (0.3%) |
| 6 | 1 (0.1%) |
| 8 | 1 (0.1%) |
| 9 | 1 (0.1%) |

| Query parameter or value | Descriptive statistics |
| --- | :---: |
| SELECT_n_of_non_aggr_func___LTRIM | [0.00, 0.00, 10.00] 0.68 / 1.14 |
| SELECT_n_of_non_aggr_func___RTRIM | |
| *0* | 606 (61%) |
| *1* | 257 (26%) |
| *10* | 1 (0.1%) |
| *2* | 80 (8.0%) |
| *3* | 28 (2.8%) |
| *4* | 17 (1.7%) |
| *5* | 7 (0.7%) |
| *6* | 3 (0.3%) |
| *7* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___UPPER | [0.00, 0.00, 9.00] 0.64 / 1.08 |
| SELECT_n_of_non_aggr_func___SQRT | |
| *0* | 780 (78%) |
| *1* | 159 (16%) |
| *2* | 30 (3.0%) |
| *3* | 12 (1.2%) |
| *4* | 11 (1.1%) |
| *5* | 2 (0.2%) |
| *6* | 5 (0.5%) |
| *7* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___SUBSTR | |
| *0* | 617 (62%) |
| *1* | 251 (25%) |
| *10* | 1 (0.1%) |
| *2* | 70 (7.0%) |
| *3* | 38 (3.8%) |
| *4* | 8 (0.8%) |
| *5* | 7 (0.7%) |
| *6* | 6 (0.6%) |
| *7* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___DAY | |
| *0* | 878 (88%) |
| *1* | 94 (9.4%) |
| *2* | 16 (1.6%) |
| *3* | 5 (0.5%) |
| *4* | 3 (0.3%) |
| *5* | 3 (0.3%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOWER | [0.00, 0.00, 9.00] 0.64 / 1.16 |
| SELECT_n_of_non_aggr_func___MONTH | [0.00, 0.00, 11.00] 0.35 / 0.98 |
| SELECT_n_of_non_aggr_func___YEAR | |
| *0* | 817 (82%) |
| *1* | 123 (12%) |
| *2* | 26 (2.6%) |
| *3* | 16 (1.6%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *4* | 8 (0.8%) |
| *5* | 7 (0.7%) |
| *6* | 1 (0.1%) |
| *7* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___TRUNC | |
| *0* | 897 (90%) |
| *1* | 87 (8.7%) |
| *2* | 9 (0.9%) |
| *3* | 4 (0.4%) |
| *4* | 1 (0.1%) |
| *5* | 1 (0.1%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___ABS | |
| *0* | 918 (92%) |
| *1* | 56 (5.6%) |
| *2* | 22 (2.2%) |
| *3* | 4 (0.4%) |
| SELECT_n_of_non_aggr_func___ROUND | |
| *0* | 907 (91%) |
| *1* | 83 (8.3%) |
| *2* | 5 (0.5%) |
| *3* | 4 (0.4%) |
| *4* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___DOW | |
| *0* | 832 (83%) |
| *1* | 117 (12%) |
| *2* | 32 (3.2%) |
| *3* | 5 (0.5%) |
| *4* | 10 (1.0%) |
| *5* | 3 (0.3%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_all_non_aggr_func | [0, 3, 83] 5 / 9 |
| SELECT_n_of_aggr_func___AVG | |
| *0* | 945 (95%) |
| *1* | 31 (3.1%) |
| *2* | 17 (1.7%) |
| *3* | 6 (0.6%) |
| *5* | 1 (0.1%) |
| SELECT_n_of_aggr_func___MAX | [0.00, 0.00, 12.00] 0.80 / 1.85 |
| SELECT_n_of_aggr_func___SUM | |
| *0* | 914 (91%) |
| *1* | 57 (5.7%) |
| *2* | 21 (2.1%) |
| *3* | 6 (0.6%) |
| *4* | 2 (0.2%) |
| SELECT_n_of_aggr_func___MIN | [0.00, 0.00, 12.00] 0.84 / 1.91 |

| Query parameter or value | Descriptive statistics |
|---|---|
| SELECT_n_of_aggr_func___COUNT | [0.00, 0.00, 12.00] 0.93 / 2.02 |
| SELECT_n_of_aggr_func___COUNT_DISTINCT | [0.00, 0.00, 13.00] 0.81 / 1.94 |
| SELECT_n_of_all_aggr_func | [0.0, 3.0, 14.0] 3.6 / 3.2 |
| FROM_n_of_join_paths | |
| *1* | 173 (17%) |
| *2* | 147 (15%) |
| *3* | 176 (18%) |
| *4* | 174 (17%) |
| *5* | 157 (16%) |
| *6* | 173 (17%) |
| FROM_n_of_super_joins___LEFT | |
| *0* | 566 (57%) |
| *1* | 296 (30%) |
| *2* | 116 (12%) |
| *3* | 19 (1.9%) |
| *4* | 3 (0.3%) |
| FROM_n_of_super_joins___RIGHT | |
| *0* | 536 (54%) |
| *1* | 336 (34%) |
| *2* | 99 (9.9%) |
| *3* | 24 (2.4%) |
| *4* | 5 (0.5%) |
| FROM_n_of_super_joins___FULL | |
| *0* | 516 (52%) |
| *1* | 338 (34%) |
| *2* | 114 (11%) |
| *3* | 27 (2.7%) |
| *4* | 5 (0.5%) |
| FROM_n_of_joins___INNER | [0.00, 3.00, 12.00] 3.33 / 2.45 |
| FROM_n_of_joins___RIGHT | [0.00, 3.00, 14.00] 3.27 / 2.58 |
| FROM_n_of_processed_rows | [5, 18,825,530, 85,814,568] 23,760,276 / 19,083,914 |
| WHERE_n_of_predicates | [0.0, 3.0, 13.0] 3.9 / 3.0 |
| WHERE_n_of_attribs_of_type___character | |
| *0* | 634 (63%) |
| *1* | 266 (27%) |
| *2* | 77 (7.7%) |
| *3* | 20 (2.0%) |
| *4* | 3 (0.3%) |
| WHERE_n_of_attribs_of_type___character_varying | |
| *0* | 343 (34%) |
| *1* | 315 (32%) |
| *2* | 190 (19%) |
| *3* | 90 (9.0%) |
| *4* | 42 (4.2%) |
| *5* | 14 (1.4%) |
| *6* | 4 (0.4%) |

| Query parameter or value | Descriptive statistics |
| --- | :---: |
| *7* | 2 (0.2%) |
| WHERE_n_of_attribs_of_type___integer | |
| *0* | 373 (37%) |
| *1* | 284 (28%) |
| *2* | 186 (19%) |
| *3* | 88 (8.8%) |
| *4* | 46 (4.6%) |
| *5* | 17 (1.7%) |
| *6* | 6 (0.6%) |
| WHERE_n_of_attribs_of_type___numeric | |
| *0* | 561 (56%) |
| *1* | 287 (29%) |
| *2* | 113 (11%) |
| *3* | 29 (2.9%) |
| *4* | 9 (0.9%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_attribs_of_type___date | |
| *0* | 771 (77%) |
| *1* | 176 (18%) |
| *2* | 45 (4.5%) |
| *3* | 4 (0.4%) |
| *4* | 3 (0.3%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_pkey_attribs | |
| *0* | 529 (53%) |
| *1* | 297 (30%) |
| *2* | 128 (13%) |
| *3* | 37 (3.7%) |
| *4* | 8 (0.8%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_connect_OR | [0.00, 1.00, 11.00] 2.27 / 2.50 |
| WHERE_n_of_operators___greater_or_less | |
| *0* | 326 (33%) |
| *1* | 290 (29%) |
| *2* | 159 (16%) |
| *3* | 111 (11%) |
| *4* | 66 (6.6%) |
| *5* | 29 (2.9%) |
| *6* | 14 (1.4%) |
| *7* | 5 (0.5%) |
| WHERE_n_of_operators___like | |
| *0* | 787 (79%) |
| *1* | 186 (19%) |
| *2* | 25 (2.5%) |
| *3* | 2 (0.2%) |
| WHERE_n_of_operators___in | |

| Query parameter or value | Descriptive statistics |
| --- | --- |
| *0* | 528 (53%) |
| *1* | 280 (28%) |
| *2* | 138 (14%) |
| *3* | 44 (4.4%) |
| *4* | 9 (0.9%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_operators___between | |
| *0* | 556 (56%) |
| *1* | 268 (27%) |
| *2* | 122 (12%) |
| *3* | 35 (3.5%) |
| *4* | 14 (1.4%) |
| *5* | 5 (0.5%) |
| WHERE_n_of_non_aggr_func___SQRT | |
| *0* | 883 (88%) |
| *1* | 112 (11%) |
| *2* | 5 (0.5%) |
| WHERE_n_of_non_aggr_func___DOW | |
| *0* | 965 (97%) |
| *1* | 33 (3.3%) |
| *2* | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___LOG | |
| *0* | 927 (93%) |
| *1* | 71 (7.1%) |
| *2* | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___YEAR | |
| *0* | 952 (95%) |
| *1* | 45 (4.5%) |
| *2* | 3 (0.3%) |
| WHERE_n_of_non_aggr_func___TRUNC | |
| *0* | 957 (96%) |
| *1* | 43 (4.3%) |
| WHERE_n_of_non_aggr_func___MONTH | |
| *0* | 963 (96%) |
| *1* | 35 (3.5%) |
| *2* | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___DAY | |
| *0* | 963 (96%) |
| *1* | 36 (3.6%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___ROUND | |
| *0* | 969 (97%) |
| *1* | 30 (3.0%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___FLOOR | |
| *0* | 963 (96%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *1* | 36 (3.6%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___ABS | |
| *0* | 960 (96%) |
| *1* | 39 (3.9%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_all_non_aggr_func | |
| *0* | 633 (63%) |
| *1* | 257 (26%) |
| *2* | 78 (7.8%) |
| *3* | 26 (2.6%) |
| *4* | 5 (0.5%) |
| *5* | 1 (0.1%) |
| GROUP_BY_n_of_columns | [0.00, 2.00, 13.00] 2.99 / 2.83 |
| HAVING_n_of_main_predicates | |
| *0* | 453 (45%) |
| *1* | 193 (19%) |
| *2* | 178 (18%) |
| *3* | 176 (18%) |
| ORDER_BY_n_of_columns | |
| *0* | 308 (31%) |
| *1* | 300 (30%) |
| *2* | 233 (23%) |
| *3* | 111 (11%) |
| *4* | 31 (3.1%) |
| *5* | 11 (1.1%) |
| *6* | 5 (0.5%) |
| *7* | 1 (0.1%) |
| limit | [2, 492, 999] 500 / 297 |
| offset | [0, 0, 1,000] 246 / 325 |