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

## Scenario: 1-1jp__ng

## Scale factor: 2 GB

## The 1000-query pack was generated on: 2023-11-08

## 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, 6, 46] 10 / 9 |
| SELECT_n_of_non_aggr_func___RTRIM | |
| 0 | 809 (81%) |
| 1 | 149 (15%) |
| 2 | 30 (3.0%) |
| 3 | 9 (0.9%) |
| 4 | 2 (0.2%) |
| 5 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___SUBSTR | |
| 0 | 803 (80%) |
| 1 | 139 (14%) |
| 2 | 50 (5.0%) |
| 3 | 3 (0.3%) |
| 4 | 4 (0.4%) |
| 5 | 1 (0.1%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| SELECT_n_of_non_aggr_func___UPPER | |
| *0* | 793 (79%) |
| *1* | 162 (16%) |
| *2* | 31 (3.1%) |
| *3* | 12 (1.2%) |
| *4* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___LTRIM | |
| *0* | 792 (79%) |
| *1* | 164 (16%) |
| *2* | 36 (3.6%) |
| *3* | 6 (0.6%) |
| *4* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___TRUNC | |
| *0* | 944 (94%) |
| *1* | 51 (5.1%) |
| *2* | 4 (0.4%) |
| *3* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOWER | |
| *0* | 789 (79%) |
| *1* | 175 (18%) |
| *2* | 27 (2.7%) |
| *3* | 7 (0.7%) |
| *4* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___SQRT | |
| *0* | 841 (84%) |
| *1* | 131 (13%) |
| *2* | 25 (2.5%) |
| *3* | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___DAY | |
| *0* | 907 (91%) |
| *1* | 80 (8.0%) |
| *2* | 13 (1.3%) |
| SELECT_n_of_non_aggr_func___MONTH | |
| *0* | 842 (84%) |
| *1* | 119 (12%) |
| *2* | 31 (3.1%) |
| *3* | 6 (0.6%) |
| *4* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___YEAR | |
| *0* | 855 (86%) |
| *1* | 102 (10%) |
| *2* | 34 (3.4%) |
| *3* | 8 (0.8%) |
| *4* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___ABS | |
| *0* | 940 (94%) |

| Query parameter or value | Descriptive statistics |
| --- | --- |
| *1* | 58 (5.8%) |
| *2* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___DOW | |
| *0* | 884 (88%) |
| *1* | 89 (8.9%) |
| *2* | 26 (2.6%) |
| *3* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___FLOOR | |
| *0* | 932 (93%) |
| *1* | 61 (6.1%) |
| *2* | 6 (0.6%) |
| *3* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOG | |
| *0* | 892 (89%) |
| *1* | 94 (9.4%) |
| *2* | 13 (1.3%) |
| *4* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___ROUND | |
| *0* | 930 (93%) |
| *1* | 67 (6.7%) |
| *2* | 3 (0.3%) |
| SELECT_n_of_all_non_aggr_func | [0.0, 1.0, 24.0] 2.5 / 3.7 |
| SELECT_n_of_aggr_func___AVG | |
| *0* | 986 (99%) |
| *1* | 12 (1.2%) |
| *2* | 2 (0.2%) |
| SELECT_n_of_aggr_func___COUNT | |
| *0* | 939 (94%) |
| *1* | 23 (2.3%) |
| *2* | 17 (1.7%) |
| *3* | 7 (0.7%) |
| *4* | 9 (0.9%) |
| *5* | 2 (0.2%) |
| *6* | 1 (0.1%) |
| *7* | 1 (0.1%) |
| *8* | 1 (0.1%) |
| SELECT_n_of_aggr_func___MIN | |
| *0* | 920 (92%) |
| *1* | 40 (4.0%) |
| *2* | 19 (1.9%) |
| *3* | 10 (1.0%) |
| *4* | 4 (0.4%) |
| *5* | 3 (0.3%) |
| *6* | 3 (0.3%) |
| *7* | 1 (0.1%) |
| SELECT_n_of_aggr_func___COUNT_DISTINCT | |

| Query parameter or value | Descriptive statistics |
|---|---|
| *0* | 936 (94%) |
| *1* | 32 (3.2%) |
| *2* | 11 (1.1%) |
| *3* | 14 (1.4%) |
| *4* | 4 (0.4%) |
| *5* | 3 (0.3%) |
| SELECT_n_of_aggr_func___SUM | |
| *0* | 991 (99%) |
| *1* | 7 (0.7%) |
| *2* | 2 (0.2%) |
| SELECT_n_of_aggr_func___MAX | |
| *0* | 927 (93%) |
| *1* | 30 (3.0%) |
| *2* | 18 (1.8%) |
| *3* | 13 (1.3%) |
| *4* | 8 (0.8%) |
| *5* | 3 (0.3%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_all_aggr_func | |
| *0* | 749 (75%) |
| *1* | 83 (8.3%) |
| *2* | 65 (6.5%) |
| *3* | 48 (4.8%) |
| *4* | 29 (2.9%) |
| *5* | 15 (1.5%) |
| *6* | 6 (0.6%) |
| *7* | 4 (0.4%) |
| *9* | 1 (0.1%) |
| FROM_n_of_join_paths | |
| *1* | 1,000 (100%) |
| FROM_n_of_joins___INNER | |
| *0* | 516 (52%) |
| *1* | 313 (31%) |
| *2* | 133 (13%) |
| *3* | 31 (3.1%) |
| *4* | 7 (0.7%) |
| FROM_n_of_joins___RIGHT | |
| *0* | 485 (49%) |
| *1* | 336 (34%) |
| *2* | 146 (15%) |
| *3* | 29 (2.9%) |
| *4* | 4 (0.4%) |
| FROM_n_of_processed_rows | [5, 1,999,865, 15,295,747] 5,711,575 / 6,221,293 |
| WHERE_n_of_predicates | |
| *0* | 224 (22%) |
| *1* | 203 (20%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *2* | 239 (24%) |
| *3* | 165 (17%) |
| *4* | 105 (11%) |
| *5* | 52 (5.2%) |
| *6* | 12 (1.2%) |
| WHERE_n_of_attribs_of_type___integer | |
| *0* | 547 (55%) |
| *1* | 339 (34%) |
| *2* | 92 (9.2%) |
| *3* | 21 (2.1%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_attribs_of_type___character_varying | |
| *0* | 525 (53%) |
| *1* | 320 (32%) |
| *2* | 134 (13%) |
| *3* | 19 (1.9%) |
| *4* | 2 (0.2%) |
| WHERE_n_of_attribs_of_type___numeric | |
| *0* | 726 (73%) |
| *1* | 223 (22%) |
| *2* | 44 (4.4%) |
| *3* | 7 (0.7%) |
| WHERE_n_of_attribs_of_type___character | |
| *0* | 808 (81%) |
| *1* | 157 (16%) |
| *2* | 35 (3.5%) |
| WHERE_n_of_attribs_of_type___date | |
| *0* | 883 (88%) |
| *1* | 110 (11%) |
| *2* | 6 (0.6%) |
| *3* | 1 (0.1%) |
| WHERE_n_of_pkey_attribs | |
| *0* | 717 (72%) |
| *1* | 242 (24%) |
| *2* | 40 (4.0%) |
| *3* | 1 (0.1%) |
| WHERE_n_of_connect_OR | |
| *0* | 580 (58%) |
| *1* | 218 (22%) |
| *2* | 116 (12%) |
| *3* | 62 (6.2%) |
| *4* | 21 (2.1%) |
| *5* | 3 (0.3%) |
| WHERE_n_of_operators___greater_or_less | |
| *0* | 489 (49%) |
| *1* | 334 (33%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *2* | 133 (13%) |
| *3* | 37 (3.7%) |
| *4* | 5 (0.5%) |
| *5* | 2 (0.2%) |
| WHERE_n_of_operators___in | |
| *0* | 704 (70%) |
| *1* | 244 (24%) |
| *2* | 49 (4.9%) |
| *3* | 2 (0.2%) |
| *4* | 1 (0.1%) |
| WHERE_n_of_operators___like | |
| *0* | 875 (88%) |
| *1* | 117 (12%) |
| *2* | 8 (0.8%) |
| WHERE_n_of_operators___between | |
| *0* | 708 (71%) |
| *1* | 245 (25%) |
| *2* | 36 (3.6%) |
| *3* | 11 (1.1%) |
| WHERE_n_of_non_aggr_func___DAY | |
| *0* | 982 (98%) |
| *1* | 18 (1.8%) |
| WHERE_n_of_non_aggr_func___TRUNC | |
| *0* | 979 (98%) |
| *1* | 18 (1.8%) |
| *2* | 3 (0.3%) |
| WHERE_n_of_non_aggr_func___SQRT | |
| *0* | 947 (95%) |
| *1* | 50 (5.0%) |
| *2* | 3 (0.3%) |
| WHERE_n_of_non_aggr_func___ABS | |
| *0* | 978 (98%) |
| *1* | 22 (2.2%) |
| WHERE_n_of_non_aggr_func___FLOOR | |
| *0* | 977 (98%) |
| *1* | 23 (2.3%) |
| WHERE_n_of_non_aggr_func___ROUND | |
| *0* | 979 (98%) |
| *1* | 21 (2.1%) |
| WHERE_n_of_non_aggr_func___DOW | |
| *0* | 981 (98%) |
| *1* | 19 (1.9%) |
| WHERE_n_of_non_aggr_func___LOG | |
| *0* | 954 (95%) |
| *1* | 45 (4.5%) |
| *2* | 1 (0.1%) |

| Query parameter or value | Descriptive statistics |
|---|:---:|
| WHERE_n_of_non_aggr_func___MONTH | |
| *0* | 985 (99%) |
| *1* | 15 (1.5%) |
| WHERE_n_of_non_aggr_func___YEAR | |
| *0* | 971 (97%) |
| *1* | 29 (2.9%) |
| WHERE_n_of_all_non_aggr_func | |
| *0* | 767 (77%) |
| *1* | 197 (20%) |
| *2* | 32 (3.2%) |
| *3* | 3 (0.3%) |
| *4* | 1 (0.1%) |
| GROUP_BY_n_of_columns | |
| *0* | 1,000 (100%) |
| HAVING_n_of_main_predicates | |
| *0* | 1,000 (100%) |
| ORDER_BY_n_of_columns | |
| *0* | 450 (45%) |
| *1* | 280 (28%) |
| *2* | 154 (15%) |
| *3* | 79 (7.9%) |
| *4* | 32 (3.2%) |
| *5* | 5 (0.5%) |
| limit | [1, 510, 1,000] 506 / 296 |
| offset | [0, 0, 999] 252 / 330 |