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

## Scenario: 1-3jp_rjjp__g__h0-3ss_h0nss

## Scale factor: 10 GB

## The 1000-query pack was generated on: 2023-11-06

## 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, 7, 91] 10 / 12 |
| SELECT_n_of_non_aggr_func___FLOOR | |
| 0 | 937 (94%) |
| 1 | 57 (5.7%) |
| 2 | 6 (0.6%) |
| SELECT_n_of_non_aggr_func___LOWER | |
| 0 | 700 (70%) |
| 1 | 221 (22%) |
| 2 | 60 (6.0%) |
| 3 | 17 (1.7%) |
| 4 | 1 (0.1%) |
| 5 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___SUBSTR | |
| 0 | 686 (69%) |
| 1 | 249 (25%) |

| Query parameter or value | Descriptive statistics |
| --- | --- |
| *2* | 47 (4.7%) |
| *3* | 12 (1.2%) |
| *4* | 5 (0.5%) |
| *6* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___RTRIM | |
| *0* | 693 (69%) |
| *1* | 222 (22%) |
| *2* | 69 (6.9%) |
| *3* | 12 (1.2%) |
| *4* | 1 (0.1%) |
| *5* | 2 (0.2%) |
| *7* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LTRIM | |
| *0* | 698 (70%) |
| *1* | 236 (24%) |
| *2* | 50 (5.0%) |
| *3* | 10 (1.0%) |
| *4* | 5 (0.5%) |
| *5* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___ABS | |
| *0* | 944 (94%) |
| *1* | 51 (5.1%) |
| *2* | 4 (0.4%) |
| *3* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOG | |
| *0* | 907 (91%) |
| *1* | 81 (8.1%) |
| *2* | 10 (1.0%) |
| *3* | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___TRUNC | |
| *0* | 944 (94%) |
| *1* | 50 (5.0%) |
| *2* | 6 (0.6%) |
| SELECT_n_of_non_aggr_func___UPPER | |
| *0* | 717 (72%) |
| *1* | 219 (22%) |
| *2* | 47 (4.7%) |
| *3* | 9 (0.9%) |
| *4* | 7 (0.7%) |
| *5* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___YEAR | |
| *0* | 879 (88%) |
| *1* | 95 (9.5%) |
| *2* | 19 (1.9%) |
| *3* | 4 (0.4%) |
| *4* | 2 (0.2%) |

| Query parameter or value | Descriptive statistics |
| --- | --- |
| *5* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___MONTH | |
| *0* | 854 (85%) |
| *1* | 113 (11%) |
| *2* | 23 (2.3%) |
| *3* | 6 (0.6%) |
| *4* | 4 (0.4%) |
| SELECT_n_of_non_aggr_func___DOW | |
| *0* | 900 (90%) |
| *1* | 74 (7.4%) |
| *2* | 20 (2.0%) |
| *3* | 5 (0.5%) |
| *4* | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___ROUND | |
| *0* | 950 (95%) |
| *1* | 45 (4.5%) |
| *2* | 5 (0.5%) |
| SELECT_n_of_non_aggr_func___DAY | |
| *0* | 924 (92%) |
| *1* | 67 (6.7%) |
| *2* | 5 (0.5%) |
| *3* | 4 (0.4%) |
| SELECT_n_of_non_aggr_func___SQRT | |
| *0* | 862 (86%) |
| *1* | 110 (11%) |
| *2* | 13 (1.3%) |
| *3* | 12 (1.2%) |
| *4* | 2 (0.2%) |
| *5* | 1 (0.1%) |
| SELECT_n_of_all_non_aggr_func | [0.0, 2.0, 40.0] 3.1 / 3.9 |
| SELECT_n_of_aggr_func___COUNT_DISTINCT | [0.00, 0.00, 9.00] 0.58 / 1.31 |
| SELECT_n_of_aggr_func___SUM | |
| *0* | 936 (94%) |
| *1* | 47 (4.7%) |
| *2* | 14 (1.4%) |
| *3* | 3 (0.3%) |
| SELECT_n_of_aggr_func___MIN | [0.00, 0.00, 10.00] 0.64 / 1.45 |
| SELECT_n_of_aggr_func___COUNT | [0.00, 0.00, 9.00] 0.67 / 1.45 |
| SELECT_n_of_aggr_func___MAX | [0.00, 0.00, 11.00] 0.62 / 1.42 |
| SELECT_n_of_aggr_func___AVG | |
| *0* | 941 (94%) |
| *1* | 46 (4.6%) |
| *2* | 10 (1.0%) |
| *3* | 3 (0.3%) |
| SELECT_n_of_all_aggr_func | [0.00, 2.00, 11.00] 2.66 / 2.17 |
| FROM_n_of_join_paths | |

| Query parameter or value | Descriptive statistics |
|---|---|
| *1* | 324 (32%) |
| *2* | 339 (34%) |
| *3* | 337 (34%) |
| FROM_n_of_super_joins___FULL | |
| *0* | 782 (78%) |
| *1* | 194 (19%) |
| *2* | 24 (2.4%) |
| FROM_n_of_super_joins___RIGHT | |
| *0* | 762 (76%) |
| *1* | 218 (22%) |
| *2* | 20 (2.0%) |
| FROM_n_of_super_joins___LEFT | |
| *0* | 764 (76%) |
| *1* | 216 (22%) |
| *2* | 20 (2.0%) |
| FROM_n_of_joins___INNER | |
| *0* | 244 (24%) |
| *1* | 264 (26%) |
| *2* | 202 (20%) |
| *3* | 149 (15%) |
| *4* | 82 (8.2%) |
| *5* | 40 (4.0%) |
| *6* | 14 (1.4%) |
| *7* | 5 (0.5%) |
| FROM_n_of_joins___RIGHT | |
| *0* | 219 (22%) |
| *1* | 291 (29%) |
| *2* | 189 (19%) |
| *3* | 156 (16%) |
| *4* | 92 (9.2%) |
| *5* | 39 (3.9%) |
| *6* | 10 (1.0%) |
| *7* | 3 (0.3%) |
| *8* | 1 (0.1%) |
| FROM_n_of_processed_rows | [5, 68,079,151, 229,441,539] 64,897,395 / 60,486,688 |
| WHERE_n_of_predicates | [0.00, 3.00, 10.00] 2.84 / 2.18 |
| WHERE_n_of_attribs_of_type___character_varying | |
| *0* | 423 (42%) |
| *1* | 319 (32%) |
| *2* | 173 (17%) |
| *3* | 72 (7.2%) |
| *4* | 10 (1.0%) |
| *5* | 2 (0.2%) |
| *6* | 1 (0.1%) |
| WHERE_n_of_attribs_of_type___character | |
| *0* | 737 (74%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *1* | 203 (20%) |
| *2* | 50 (5.0%) |
| *3* | 9 (0.9%) |
| *4* | 1 (0.1%) |
| WHERE_n_of_attribs_of_type___integer | |
| *0* | 442 (44%) |
| *1* | 312 (31%) |
| *2* | 163 (16%) |
| *3* | 60 (6.0%) |
| *4* | 19 (1.9%) |
| *5* | 4 (0.4%) |
| WHERE_n_of_attribs_of_type___numeric | |
| *0* | 652 (65%) |
| *1* | 249 (25%) |
| *2* | 77 (7.7%) |
| *3* | 21 (2.1%) |
| *4* | 1 (0.1%) |
| WHERE_n_of_attribs_of_type___date | |
| *0* | 839 (84%) |
| *1* | 142 (14%) |
| *2* | 16 (1.6%) |
| *3* | 3 (0.3%) |
| WHERE_n_of_pkey_attribs | |
| *0* | 615 (62%) |
| *1* | 284 (28%) |
| *2* | 83 (8.3%) |
| *3* | 16 (1.6%) |
| *4* | 2 (0.2%) |
| WHERE_n_of_connect_OR | [0.00, 1.00, 9.00] 1.36 / 1.68 |
| WHERE_n_of_operators___greater_or_less | |
| *0* | 389 (39%) |
| *1* | 298 (30%) |
| *2* | 193 (19%) |
| *3* | 78 (7.8%) |
| *4* | 32 (3.2%) |
| *5* | 8 (0.8%) |
| *6* | 2 (0.2%) |
| WHERE_n_of_operators___between | |
| *0* | 599 (60%) |
| *1* | 294 (29%) |
| *2* | 84 (8.4%) |
| *3* | 18 (1.8%) |
| *4* | 4 (0.4%) |
| *5* | 1 (0.1%) |
| WHERE_n_of_operators___in | |
| *0* | 636 (64%) |

| Query parameter or value | Descriptive statistics |
| --- | --- |
| *1* | 259 (26%) |
| *2* | 79 (7.9%) |
| *3* | 24 (2.4%) |
| *4* | 2 (0.2%) |
| WHERE_n_of_operators___like | |
| *0* | 841 (84%) |
| *1* | 143 (14%) |
| *2* | 13 (1.3%) |
| *3* | 3 (0.3%) |
| WHERE_n_of_non_aggr_func___TRUNC | |
| *0* | 972 (97%) |
| *1* | 28 (2.8%) |
| WHERE_n_of_non_aggr_func___SQRT | |
| *0* | 912 (91%) |
| *1* | 84 (8.4%) |
| *2* | 3 (0.3%) |
| *3* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___ABS | |
| *0* | 971 (97%) |
| *1* | 29 (2.9%) |
| WHERE_n_of_non_aggr_func___DAY | |
| *0* | 975 (98%) |
| *1* | 25 (2.5%) |
| WHERE_n_of_non_aggr_func___DOW | |
| *0* | 977 (98%) |
| *1* | 22 (2.2%) |
| *2* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___MONTH | |
| *0* | 971 (97%) |
| *1* | 29 (2.9%) |
| WHERE_n_of_non_aggr_func___LOG | |
| *0* | 949 (95%) |
| *1* | 46 (4.6%) |
| *2* | 4 (0.4%) |
| *3* | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___YEAR | |
| *0* | 977 (98%) |
| *1* | 23 (2.3%) |
| WHERE_n_of_non_aggr_func___ROUND | |
| *0* | 966 (97%) |
| *1* | 34 (3.4%) |
| WHERE_n_of_non_aggr_func___FLOOR | |
| *0* | 977 (98%) |
| *1* | 23 (2.3%) |
| WHERE_n_of_all_non_aggr_func | |
| *0* | 719 (72%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| *1* | 214 (21%) |
| *2* | 52 (5.2%) |
| *3* | 13 (1.3%) |
| *4* | 2 (0.2%) |
| GROUP_BY_n_of_columns | [0.00, 2.00, 9.00] 2.44 / 2.17 |
| HAVING_n_of_main_predicates | |
| *0* | 287 (29%) |
| *1* | 91 (9.1%) |
| *2* | 153 (15%) |
| *3* | 186 (19%) |
| *4* | 157 (16%) |
| *5* | 85 (8.5%) |
| *6* | 41 (4.1%) |
| HAVING_n_of_main_predicates___scalar_subquery | |
| *0* | 450 (45%) |
| *1* | 173 (17%) |
| *2* | 177 (18%) |
| *3* | 200 (20%) |
| HAVING_n_of_subqueries___scalar_subquery | [0.0, 1.0, 27.0] 2.4 / 3.6 |
| HAVING_n_of_processed_rows_by_subqueries | [1,900,338, 3,913,916,031] 147,505,392 / 325,285,514 |
| ORDER_BY_n_of_columns | |
| *0* | 313 (31%) |
| *1* | 378 (38%) |
| *2* | 220 (22%) |
| *3* | 72 (7.2%) |
| *4* | 10 (1.0%) |
| *5* | 6 (0.6%) |
| *6* | 1 (0.1%) |
| limit | [1, 484, 999] 495 / 290 |
| offset | [0, 15, 999] 254 / 320 |