

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

Scenario: 1-3jp_rjpp__g__h0-3ss_h0-3nss

Scale factor: 1000 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, 6, 101] 10 / 13 |
| SELECT_n_of_non_aggr_func__LOWER | |
| 0 | 699 (70%) |
| 1 | 229 (23%) |
| 2 | 56 (5.6%) |
| 3 | 13 (1.3%) |
| 4 | 3 (0.3%) |
| SELECT_n_of_non_aggr_func__UPPER | |
| 0 | 671 (67%) |
| 1 | 266 (27%) |
| 2 | 37 (3.7%) |
| 3 | 17 (1.7%) |
| 4 | 4 (0.4%) |
| 5 | 4 (0.4%) |
| 6 | 1 (0.1%) |

| Query parameter or value | Descriptive statistics |
|---------------------------------------|------------------------|
| SELECT__n_of__non__aggr__func__DOW | |
| 0 | 871 (87%) |
| 1 | 107 (11%) |
| 2 | 18 (1.8%) |
| 3 | 3 (0.3%) |
| 4 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__MONTH | |
| 0 | 857 (86%) |
| 1 | 107 (11%) |
| 2 | 22 (2.2%) |
| 3 | 8 (0.8%) |
| 4 | 3 (0.3%) |
| 5 | 1 (0.1%) |
| 6 | 1 (0.1%) |
| 7 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__RTRIM | |
| 0 | 670 (67%) |
| 1 | 258 (26%) |
| 2 | 47 (4.7%) |
| 3 | 19 (1.9%) |
| 4 | 4 (0.4%) |
| 5 | 2 (0.2%) |
| SELECT__n_of__non__aggr__func__SUBSTR | |
| 0 | 690 (69%) |
| 1 | 233 (23%) |
| 2 | 60 (6.0%) |
| 3 | 12 (1.2%) |
| 4 | 3 (0.3%) |
| 5 | 1 (0.1%) |
| 6 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__YEAR | |
| 0 | 872 (87%) |
| 1 | 102 (10%) |
| 2 | 19 (1.9%) |
| 3 | 4 (0.4%) |
| 4 | 2 (0.2%) |
| 5 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__LTRIM | |
| 0 | 664 (66%) |
| 1 | 259 (26%) |
| 2 | 60 (6.0%) |
| 3 | 14 (1.4%) |
| 4 | 1 (0.1%) |
| 5 | 2 (0.2%) |
| SELECT__n_of__non__aggr__func__ABS | |
| 0 | 940 (94%) |

| Query parameter or value | Descriptive statistics |
|-----------------------------------|---------------------------------|
| 1 | 54 (5.4%) |
| 2 | 6 (0.6%) |
| SELECT_n_of_non_aggr_func___SQRT | |
| 0 | 835 (84%) |
| 1 | 126 (13%) |
| 2 | 28 (2.8%) |
| 3 | 9 (0.9%) |
| 4 | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___DAY | |
| 0 | 917 (92%) |
| 1 | 69 (6.9%) |
| 2 | 10 (1.0%) |
| 3 | 2 (0.2%) |
| 4 | 1 (0.1%) |
| 5 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOG | |
| 0 | 876 (88%) |
| 1 | 104 (10%) |
| 2 | 17 (1.7%) |
| 3 | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___TRUNC | |
| 0 | 941 (94%) |
| 1 | 55 (5.5%) |
| 2 | 3 (0.3%) |
| 3 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___FLOOR | |
| 0 | 945 (95%) |
| 1 | 48 (4.8%) |
| 2 | 7 (0.7%) |
| SELECT_n_of_non_aggr_func___ROUND | |
| 0 | 933 (93%) |
| 1 | 65 (6.5%) |
| 2 | 2 (0.2%) |
| SELECT_n_of_all_non_aggr_func | [0.0, 2.0, 34.0] 3.4 / 4.3 |
| SELECT_n_of_aggr_func___MAX | [0.00, 0.00, 11.00] 0.61 / 1.36 |
| SELECT_n_of_aggr_func___MIN | [0.00, 0.00, 9.00] 0.64 / 1.44 |
| SELECT_n_of_aggr_func___SUM | |
| 0 | 948 (95%) |
| 1 | 34 (3.4%) |
| 2 | 10 (1.0%) |
| 3 | 7 (0.7%) |
| 4 | 1 (0.1%) |
| SELECT_n_of_aggr_func___AVG | |
| 0 | 951 (95%) |
| 1 | 30 (3.0%) |
| 2 | 16 (1.6%) |

| Query parameter or value | Descriptive statistics |
|---------------------------------------|--|
| 3 | 2 (0.2%) |
| 4 | 1 (0.1%) |
| SELECT_n_of_aggr_func__COUNT | [0.00, 0.00, 10.00] 0.63 / 1.49 |
| SELECT_n_of_aggr_func__COUNT_DISTINCT | [0.00, 0.00, 10.00] 0.66 / 1.54 |
| SELECT_n_of_all_aggr_func | [0.00, 2.00, 11.00] 2.70 / 2.33 |
| FROM_n_of_join_paths | |
| 1 | 312 (31%) |
| 2 | 325 (33%) |
| 3 | 363 (36%) |
| FROM_n_of_super_joins__RIGHT | |
| 0 | 749 (75%) |
| 1 | 235 (24%) |
| 2 | 16 (1.6%) |
| FROM_n_of_super_joins__LEFT | |
| 0 | 754 (75%) |
| 1 | 225 (23%) |
| 2 | 21 (2.1%) |
| FROM_n_of_super_joins__FULL | |
| 0 | 778 (78%) |
| 1 | 206 (21%) |
| 2 | 16 (1.6%) |
| FROM_n_of_joins__INNER | |
| 0 | 228 (23%) |
| 1 | 269 (27%) |
| 2 | 217 (22%) |
| 3 | 145 (15%) |
| 4 | 89 (8.9%) |
| 5 | 36 (3.6%) |
| 6 | 11 (1.1%) |
| 7 | 5 (0.5%) |
| FROM_n_of_joins__RIGHT | |
| 0 | 240 (24%) |
| 1 | 236 (24%) |
| 2 | 219 (22%) |
| 3 | 156 (16%) |
| 4 | 92 (9.2%) |
| 5 | 36 (3.6%) |
| 6 | 12 (1.2%) |
| 7 | 9 (0.9%) |
| FROM_n_of_processed_rows | [5, 6,819,989,749, 22,299,969,187] 6,714,651,347 / 5,980,482,463 |
| WHERE_n_of_predicates | [0.00, 2.00, 9.00] 2.63 / 2.17 |
| WHERE_n_of_attribs_of_type__character | |
| 0 | 740 (74%) |
| 1 | 209 (21%) |
| 2 | 41 (4.1%) |

| Query parameter or value | Descriptive statistics |
|---|------------------------|
| 3 | 8 (0.8%) |
| 4 | 2 (0.2%) |
| WHERE_n_of_attribs_of_type__character_varying | |
| 0 | 490 (49%) |
| 1 | 298 (30%) |
| 2 | 134 (13%) |
| 3 | 54 (5.4%) |
| 4 | 20 (2.0%) |
| 5 | 2 (0.2%) |
| 6 | 2 (0.2%) |
| WHERE_n_of_attribs_of_type__integer | |
| 0 | 473 (47%) |
| 1 | 303 (30%) |
| 2 | 159 (16%) |
| 3 | 55 (5.5%) |
| 4 | 10 (1.0%) |
| WHERE_n_of_attribs_of_type__numeric | |
| 0 | 630 (63%) |
| 1 | 286 (29%) |
| 2 | 65 (6.5%) |
| 3 | 16 (1.6%) |
| 4 | 3 (0.3%) |
| WHERE_n_of_attribs_of_type__date | |
| 0 | 841 (84%) |
| 1 | 145 (15%) |
| 2 | 9 (0.9%) |
| 3 | 5 (0.5%) |
| WHERE_n_of_pkey_attribs | |
| 0 | 651 (65%) |
| 1 | 256 (26%) |
| 2 | 81 (8.1%) |
| 3 | 11 (1.1%) |
| 4 | 1 (0.1%) |
| WHERE_n_of_connect_OR | |
| 0 | 488 (49%) |
| 1 | 178 (18%) |
| 2 | 124 (12%) |
| 3 | 96 (9.6%) |
| 4 | 52 (5.2%) |
| 5 | 37 (3.7%) |
| 6 | 20 (2.0%) |
| 7 | 5 (0.5%) |
| WHERE_n_of_operators__in | |
| 0 | 639 (64%) |
| 1 | 267 (27%) |
| 2 | 77 (7.7%) |

| Query parameter or value | Descriptive statistics |
|---------------------------------------|------------------------|
| 3 | 15 (1.5%) |
| 4 | 2 (0.2%) |
| WHERE_n_of_operators__like | |
| 0 | 845 (85%) |
| 1 | 139 (14%) |
| 2 | 14 (1.4%) |
| 3 | 2 (0.2%) |
| WHERE_n_of_operators__between | |
| 0 | 629 (63%) |
| 1 | 275 (28%) |
| 2 | 82 (8.2%) |
| 3 | 13 (1.3%) |
| 4 | 1 (0.1%) |
| WHERE_n_of_operators__greater_or_less | |
| 0 | 437 (44%) |
| 1 | 276 (28%) |
| 2 | 158 (16%) |
| 3 | 96 (9.6%) |
| 4 | 25 (2.5%) |
| 5 | 7 (0.7%) |
| 6 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func__ROUND | |
| 0 | 978 (98%) |
| 1 | 22 (2.2%) |
| WHERE_n_of_non_aggr_func__SQRT | |
| 0 | 895 (90%) |
| 1 | 98 (9.8%) |
| 2 | 7 (0.7%) |
| WHERE_n_of_non_aggr_func__FLOOR | |
| 0 | 976 (98%) |
| 1 | 22 (2.2%) |
| 2 | 2 (0.2%) |
| WHERE_n_of_non_aggr_func__LOG | |
| 0 | 932 (93%) |
| 1 | 64 (6.4%) |
| 2 | 4 (0.4%) |
| WHERE_n_of_non_aggr_func__DAY | |
| 0 | 977 (98%) |
| 1 | 23 (2.3%) |
| WHERE_n_of_non_aggr_func__TRUNC | |
| 0 | 974 (97%) |
| 1 | 26 (2.6%) |
| WHERE_n_of_non_aggr_func__DOW | |
| 0 | 972 (97%) |
| 1 | 27 (2.7%) |
| 2 | 1 (0.1%) |

| Query parameter or value | Descriptive statistics |
|--|---|
| WHERE_n_of_non_aggr_func__YEAR | |
| 0 | 971 (97%) |
| 1 | 29 (2.9%) |
| WHERE_n_of_non_aggr_func__ABS | |
| 0 | 975 (98%) |
| 1 | 24 (2.4%) |
| 2 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func__MONTH | |
| 0 | 970 (97%) |
| 1 | 30 (3.0%) |
| WHERE_n_of_all_non_aggr_func | |
| 0 | 701 (70%) |
| 1 | 221 (22%) |
| 2 | 65 (6.5%) |
| 3 | 10 (1.0%) |
| 4 | 1 (0.1%) |
| 5 | 2 (0.2%) |
| GROUP_BY_n_of_columns | [0.00, 2.00, 9.00] 2.28 / 2.08 |
| HAVING_n_of_main_predicates | [0.00, 4.00, 9.00] 3.43 / 2.63 |
| HAVING_n_of_main_predicates__scalar_subquery | |
| 0 | 438 (44%) |
| 1 | 182 (18%) |
| 2 | 177 (18%) |
| 3 | 203 (20%) |
| HAVING_n_of_main_predicates__non_scalar_subquery | |
| 0 | 436 (44%) |
| 1 | 194 (19%) |
| 2 | 170 (17%) |
| 3 | 200 (20%) |
| HAVING_n_of_subqueries__scalar_subquery | [0.0, 1.0, 20.0] 2.5 / 3.8 |
| HAVING_n_of_subqueries__non_scalar_subquery | |
| 0 | 436 (44%) |
| 1 | 194 (19%) |
| 2 | 170 (17%) |
| 3 | 200 (20%) |
| HAVING_n_of_processed_rows_by_subqueries | [0, 10, 100,000,000, 395,279,445,186] 24,455,598,568 / 44,390,056,141 |
| ORDER_BY_n_of_columns | |
| 0 | 261 (26%) |
| 1 | 379 (38%) |
| 2 | 266 (27%) |
| 3 | 76 (7.6%) |
| 4 | 10 (1.0%) |
| 5 | 7 (0.7%) |
| 7 | 1 (0.1%) |
| limit | [1, 480, 1,000] 493 / 287 |

| Query parameter or value | Descriptive statistics |
|--------------------------|------------------------|
| offset | [0, 56, 996] 255 / 320 |