

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

Scenario: 1-6jp_ujpp__g__h0-3ss_h0-3nss

Scale factor: 50 GB

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

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, 187] 17 / 27 |
| SELECT_n_of_non_aggr_func__LTRIM | |
| 0 | 596 (60%) |
| 1 | 275 (28%) |
| 13 | 1 (0.1%) |
| 2 | 83 (8.3%) |
| 3 | 17 (1.7%) |
| 4 | 15 (1.5%) |
| 5 | 7 (0.7%) |
| 6 | 5 (0.5%) |
| 7 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func__MONTH | [0.00, 0.00, 10.00] 0.35 / 0.98 |
| SELECT_n_of_non_aggr_func__YEAR | [0.00, 0.00, 10.00] 0.29 / 0.93 |
| SELECT_n_of_non_aggr_func__ABS | |
| 0 | 909 (91%) |

| Query parameter or value | Descriptive statistics |
|------------------------------------|---------------------------------|
| 1 | 71 (7.1%) |
| 2 | 13 (1.3%) |
| 3 | 3 (0.3%) |
| 4 | 4 (0.4%) |
| SELECT_n_of_non_aggr_func___DOW | |
| 0 | 863 (86%) |
| 1 | 94 (9.4%) |
| 2 | 21 (2.1%) |
| 3 | 13 (1.3%) |
| 4 | 3 (0.3%) |
| 5 | 3 (0.3%) |
| 6 | 2 (0.2%) |
| 7 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___FLOOR | |
| 0 | 906 (91%) |
| 1 | 77 (7.7%) |
| 2 | 14 (1.4%) |
| 3 | 3 (0.3%) |
| SELECT_n_of_non_aggr_func___LOG | |
| 0 | 867 (87%) |
| 1 | 94 (9.4%) |
| 2 | 20 (2.0%) |
| 3 | 7 (0.7%) |
| 4 | 8 (0.8%) |
| 5 | 2 (0.2%) |
| 6 | 1 (0.1%) |
| 8 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___LOWER | [0.00, 0.00, 9.00] 0.63 / 1.14 |
| SELECT_n_of_non_aggr_func___RTRIM | |
| 0 | 644 (64%) |
| 1 | 253 (25%) |
| 2 | 65 (6.5%) |
| 3 | 17 (1.7%) |
| 4 | 9 (0.9%) |
| 5 | 8 (0.8%) |
| 7 | 2 (0.2%) |
| 8 | 2 (0.2%) |
| SELECT_n_of_non_aggr_func___SUBSTR | [0.00, 0.00, 9.00] 0.63 / 1.09 |
| SELECT_n_of_non_aggr_func___TRUNC | |
| 0 | 908 (91%) |
| 1 | 77 (7.7%) |
| 2 | 10 (1.0%) |
| 3 | 5 (0.5%) |
| SELECT_n_of_non_aggr_func___UPPER | [0.00, 0.00, 11.00] 0.57 / 1.05 |
| SELECT_n_of_non_aggr_func___DAY | |
| 0 | 880 (88%) |

| Query parameter or value | Descriptive statistics |
|--|---------------------------------|
| 1 | 86 (8.6%) |
| 2 | 23 (2.3%) |
| 3 | 8 (0.8%) |
| 4 | 2 (0.2%) |
| 5 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___SQRT | |
| 0 | 785 (79%) |
| 1 | 143 (14%) |
| 2 | 37 (3.7%) |
| 3 | 13 (1.3%) |
| 4 | 8 (0.8%) |
| 5 | 8 (0.8%) |
| 6 | 4 (0.4%) |
| 8 | 1 (0.1%) |
| 9 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func___ROUND | |
| 0 | 915 (92%) |
| 1 | 72 (7.2%) |
| 2 | 8 (0.8%) |
| 3 | 5 (0.5%) |
| SELECT_n_of_all_non_aggr_func | [0, 3, 81] 5 / 8 |
| SELECT_n_of_aggr_func___MAX | [0.00, 0.00, 13.00] 0.83 / 1.85 |
| SELECT_n_of_aggr_func___COUNT | [0.00, 0.00, 11.00] 0.81 / 1.87 |
| SELECT_n_of_aggr_func___MIN | [0.00, 0.00, 12.00] 0.72 / 1.72 |
| SELECT_n_of_aggr_func___COUNT_DISTINCT | [0.00, 0.00, 12.00] 0.89 / 1.95 |
| SELECT_n_of_aggr_func___AVG | |
| 0 | 940 (94%) |
| 1 | 41 (4.1%) |
| 2 | 10 (1.0%) |
| 3 | 7 (0.7%) |
| 4 | 1 (0.1%) |
| 5 | 1 (0.1%) |
| SELECT_n_of_aggr_func___SUM | |
| 0 | 937 (94%) |
| 1 | 36 (3.6%) |
| 2 | 20 (2.0%) |
| 3 | 6 (0.6%) |
| 4 | 1 (0.1%) |
| SELECT_n_of_all_aggr_func | [0.00, 3.00, 14.00] 3.44 / 2.99 |
| FROM_n_of_join_paths | |
| 1 | 168 (17%) |
| 2 | 152 (15%) |
| 3 | 143 (14%) |
| 4 | 186 (19%) |
| 5 | 173 (17%) |
| 6 | 178 (18%) |

| Query parameter or value | Descriptive statistics |
|---|---|
| FROM_n_of_super_joins__RIGHT | |
| 0 | 551 (55%) |
| 1 | 301 (30%) |
| 2 | 122 (12%) |
| 3 | 19 (1.9%) |
| 4 | 7 (0.7%) |
| FROM_n_of_super_joins__LEFT | |
| 0 | 530 (53%) |
| 1 | 305 (31%) |
| 2 | 129 (13%) |
| 3 | 33 (3.3%) |
| 4 | 3 (0.3%) |
| FROM_n_of_super_joins__FULL | |
| 0 | 553 (55%) |
| 1 | 306 (31%) |
| 2 | 107 (11%) |
| 3 | 33 (3.3%) |
| 4 | 1 (0.1%) |
| FROM_n_of_joins__INNER | [0.00, 3.00, 14.00] 3.40 / 2.53 |
| FROM_n_of_joins__RIGHT | [0.00, 3.00, 15.00] 3.31 / 2.43 |
| FROM_n_of_processed_rows | [5, 468,264,244, 2,211,049,523] 612,991,169 / 500,996,160 |
| WHERE_n_of_predicates | [0.00, 4.00, 12.00] 3.95 / 2.99 |
| WHERE_n_of_attribs_of_type__integer | |
| 0 | 366 (37%) |
| 1 | 302 (30%) |
| 2 | 170 (17%) |
| 3 | 98 (9.8%) |
| 4 | 42 (4.2%) |
| 5 | 17 (1.7%) |
| 6 | 1 (0.1%) |
| 7 | 4 (0.4%) |
| WHERE_n_of_attribs_of_type__character_varying | |
| 0 | 347 (35%) |
| 1 | 303 (30%) |
| 2 | 195 (20%) |
| 3 | 99 (9.9%) |
| 4 | 29 (2.9%) |
| 5 | 20 (2.0%) |
| 6 | 7 (0.7%) |
| WHERE_n_of_attribs_of_type__date | |
| 0 | 771 (77%) |
| 1 | 177 (18%) |
| 2 | 44 (4.4%) |
| 3 | 6 (0.6%) |
| 4 | 2 (0.2%) |

| Query parameter or value | Descriptive statistics |
|---------------------------------------|---------------------------------|
| WHERE_n_of_attribs_of_type__character | |
| 0 | 638 (64%) |
| 1 | 232 (23%) |
| 2 | 97 (9.7%) |
| 3 | 25 (2.5%) |
| 4 | 6 (0.6%) |
| 5 | 2 (0.2%) |
| WHERE_n_of_attribs_of_type__numeric | |
| 0 | 571 (57%) |
| 1 | 267 (27%) |
| 2 | 115 (12%) |
| 3 | 34 (3.4%) |
| 4 | 10 (1.0%) |
| 5 | 2 (0.2%) |
| 6 | 1 (0.1%) |
| WHERE_n_of_pkey_attribs | |
| 0 | 540 (54%) |
| 1 | 294 (29%) |
| 2 | 112 (11%) |
| 3 | 39 (3.9%) |
| 4 | 10 (1.0%) |
| 5 | 3 (0.3%) |
| 6 | 1 (0.1%) |
| 7 | 1 (0.1%) |
| WHERE_n_of_connect_OR | [0.00, 2.00, 11.00] 2.31 / 2.49 |
| WHERE_n_of_operators__between | |
| 0 | 505 (51%) |
| 1 | 280 (28%) |
| 2 | 150 (15%) |
| 3 | 43 (4.3%) |
| 4 | 17 (1.7%) |
| 5 | 2 (0.2%) |
| 6 | 3 (0.3%) |
| WHERE_n_of_operators__greater_or_less | |
| 0 | 318 (32%) |
| 1 | 259 (26%) |
| 2 | 200 (20%) |
| 3 | 128 (13%) |
| 4 | 59 (5.9%) |
| 5 | 20 (2.0%) |
| 6 | 11 (1.1%) |
| 7 | 4 (0.4%) |
| 8 | 1 (0.1%) |
| WHERE_n_of_operators__in | |
| 0 | 545 (55%) |
| 1 | 276 (28%) |

| Query parameter or value | Descriptive statistics |
|---------------------------------|------------------------|
| 2 | 121 (12%) |
| 3 | 40 (4.0%) |
| 4 | 12 (1.2%) |
| 5 | 5 (0.5%) |
| 6 | 1 (0.1%) |
| WHERE_n_of_operators__like | |
| 0 | 792 (79%) |
| 1 | 182 (18%) |
| 2 | 22 (2.2%) |
| 3 | 4 (0.4%) |
| WHERE_n_of_non_aggr_func__MONTH | |
| 0 | 958 (96%) |
| 1 | 40 (4.0%) |
| 2 | 2 (0.2%) |
| WHERE_n_of_non_aggr_func__LOG | |
| 0 | 918 (92%) |
| 1 | 76 (7.6%) |
| 2 | 6 (0.6%) |
| WHERE_n_of_non_aggr_func__SQRT | |
| 0 | 870 (87%) |
| 1 | 120 (12%) |
| 2 | 9 (0.9%) |
| 3 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func__YEAR | |
| 0 | 956 (96%) |
| 1 | 44 (4.4%) |
| WHERE_n_of_non_aggr_func__DOW | |
| 0 | 957 (96%) |
| 1 | 42 (4.2%) |
| 2 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func__DAY | |
| 0 | 960 (96%) |
| 1 | 40 (4.0%) |
| WHERE_n_of_non_aggr_func__FLOOR | |
| 0 | 960 (96%) |
| 1 | 39 (3.9%) |
| 2 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func__TRUNC | |
| 0 | 961 (96%) |
| 1 | 38 (3.8%) |
| 3 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func__ABS | |
| 0 | 963 (96%) |
| 1 | 35 (3.5%) |
| 2 | 2 (0.2%) |
| WHERE_n_of_non_aggr_func__ROUND | |

| Query parameter or value | Descriptive statistics |
|--|--|
| 0 | 955 (96%) |
| 1 | 43 (4.3%) |
| 2 | 2 (0.2%) |
| WHERE_n_of_all_non_aggr_func | |
| 0 | 605 (61%) |
| 1 | 270 (27%) |
| 2 | 88 (8.8%) |
| 3 | 29 (2.9%) |
| 4 | 5 (0.5%) |
| 5 | 2 (0.2%) |
| 6 | 1 (0.1%) |
| GROUP_BY_n_of_columns | [0.00, 2.00, 11.00] 3.10 / 2.92 |
| HAVING_n_of_main_predicates | [0.00, 4.00, 9.00] 3.46 / 2.64 |
| HAVING_n_of_main_predicates__non_scalar_subquery | |
| 0 | 421 (42%) |
| 1 | 182 (18%) |
| 2 | 183 (18%) |
| 3 | 214 (21%) |
| HAVING_n_of_main_predicates__scalar_subquery | |
| 0 | 429 (43%) |
| 1 | 190 (19%) |
| 2 | 183 (18%) |
| 3 | 198 (20%) |
| HAVING_n_of_subqueries__non_scalar_subquery | |
| 0 | 421 (42%) |
| 1 | 182 (18%) |
| 2 | 183 (18%) |
| 3 | 214 (21%) |
| HAVING_n_of_subqueries__scalar_subquery | [0.0, 1.0, 25.0] 2.4 / 3.6 |
| HAVING_n_of_processed_rows_by_subqueries | [0, 685,660,331, 28,229,847,120] 2,179,069,991 / 3,619,516,928 |
| ORDER_BY_n_of_columns | |
| 0 | 303 (30%) |
| 1 | 338 (34%) |
| 2 | 228 (23%) |
| 3 | 87 (8.7%) |
| 4 | 25 (2.5%) |
| 5 | 13 (1.3%) |
| 6 | 1 (0.1%) |
| 7 | 2 (0.2%) |
| 8 | 3 (0.3%) |
| limit | [1, 515, 1,000] 509 / 292 |
| offset | [0, 0, 999] 243 / 321 |