

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

Scenario: 1-6jp_ujpp__g__h0-3ss_h0nss

Scale factor: 0.01 GB

The 1000-query pack was generated on: 2023-11-10

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, 183] 16 / 26 |
| SELECT_n_of_non_aggr_func__LOWER | [0.00, 0.00, 15.00] 0.66 / 1.15 |
| SELECT_n_of_non_aggr_func__SQRT | |
| 0 | 798 (80%) |
| 1 | 144 (14%) |
| 2 | 29 (2.9%) |
| 3 | 14 (1.4%) |
| 4 | 9 (0.9%) |
| 5 | 3 (0.3%) |
| 6 | 1 (0.1%) |
| 7 | 2 (0.2%) |
| SELECT_n_of_non_aggr_func__UPPER | [0.00, 0.00, 11.00] 0.58 / 1.02 |
| SELECT_n_of_non_aggr_func__LTRIM | |
| 0 | 627 (63%) |
| 1 | 265 (27%) |

| Query parameter or value | Descriptive statistics |
|-----------------------------------|------------------------|
| 2 | 67 (6.7%) |
| 3 | 18 (1.8%) |
| 4 | 12 (1.2%) |
| 5 | 5 (0.5%) |
| 6 | 4 (0.4%) |
| 7 | 1 (0.1%) |
| 9 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func__LOG | |
| 0 | 887 (89%) |
| 1 | 79 (7.9%) |
| 2 | 21 (2.1%) |
| 3 | 7 (0.7%) |
| 4 | 4 (0.4%) |
| 5 | 1 (0.1%) |
| 6 | 1 (0.1%) |
| SELECT_n_of_non_aggr_func__SUBSTR | |
| 0 | 617 (62%) |
| 1 | 268 (27%) |
| 2 | 64 (6.4%) |
| 3 | 27 (2.7%) |
| 4 | 12 (1.2%) |
| 5 | 5 (0.5%) |
| 6 | 1 (0.1%) |
| 7 | 3 (0.3%) |
| 8 | 3 (0.3%) |
| SELECT_n_of_non_aggr_func__RTRIM | |
| 0 | 627 (63%) |
| 1 | 254 (25%) |
| 2 | 76 (7.6%) |
| 3 | 20 (2.0%) |
| 4 | 9 (0.9%) |
| 5 | 8 (0.8%) |
| 6 | 2 (0.2%) |
| 7 | 2 (0.2%) |
| 8 | 2 (0.2%) |
| SELECT_n_of_non_aggr_func__TRUNC | |
| 0 | 916 (92%) |
| 1 | 74 (7.4%) |
| 2 | 8 (0.8%) |
| 3 | 2 (0.2%) |
| SELECT_n_of_non_aggr_func__ROUND | |
| 0 | 920 (92%) |
| 1 | 65 (6.5%) |
| 2 | 10 (1.0%) |
| 3 | 3 (0.3%) |
| 4 | 2 (0.2%) |

| Query parameter or value | Descriptive statistics |
|--------------------------------------|---------------------------------|
| SELECT__n_of__non__aggr__func__ABS | |
| 0 | 926 (93%) |
| 1 | 59 (5.9%) |
| 2 | 12 (1.2%) |
| 3 | 2 (0.2%) |
| 4 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__YEAR | |
| 0 | 842 (84%) |
| 1 | 107 (11%) |
| 2 | 26 (2.6%) |
| 3 | 12 (1.2%) |
| 4 | 3 (0.3%) |
| 5 | 6 (0.6%) |
| 6 | 2 (0.2%) |
| 7 | 1 (0.1%) |
| 8 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__FLOOR | |
| 0 | 924 (92%) |
| 1 | 61 (6.1%) |
| 2 | 12 (1.2%) |
| 3 | 2 (0.2%) |
| 4 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__MONTH | [0.00, 0.00, 11.00] 0.34 / 0.95 |
| SELECT__n_of__non__aggr__func__DAY | |
| 0 | 880 (88%) |
| 1 | 89 (8.9%) |
| 2 | 20 (2.0%) |
| 3 | 5 (0.5%) |
| 4 | 3 (0.3%) |
| 5 | 1 (0.1%) |
| 6 | 1 (0.1%) |
| 8 | 1 (0.1%) |
| SELECT__n_of__non__aggr__func__DOW | |
| 0 | 861 (86%) |
| 1 | 101 (10%) |
| 2 | 20 (2.0%) |
| 3 | 7 (0.7%) |
| 4 | 6 (0.6%) |
| 5 | 1 (0.1%) |
| 6 | 3 (0.3%) |
| 7 | 1 (0.1%) |
| SELECT__n_of__all__non__aggr__func | [0.0, 3.0, 81.0] 4.8 / 8.0 |
| SELECT__n_of__aggr__func__COUNT | [0.00, 0.00, 13.00] 0.87 / 2.03 |
| SELECT__n_of__aggr__func__MAX | [0.00, 0.00, 13.00] 0.96 / 2.05 |
| SELECT__n_of__aggr__func__SUM | |
| 0 | 937 (94%) |

| Query parameter or value | Descriptive statistics |
|--|---------------------------------------|
| 1 | 40 (4.0%) |
| 2 | 17 (1.7%) |
| 3 | 4 (0.4%) |
| 4 | 2 (0.2%) |
| SELECT_n_of_aggr_func__MIN | [0.00, 0.00, 13.00] 0.89 / 2.04 |
| SELECT_n_of_aggr_func__COUNT_DISTINCT | [0.00, 0.00, 12.00] 0.73 / 1.69 |
| SELECT_n_of_aggr_func__AVG | |
| 0 | 936 (94%) |
| 1 | 46 (4.6%) |
| 2 | 14 (1.4%) |
| 3 | 1 (0.1%) |
| 4 | 2 (0.2%) |
| 5 | 1 (0.1%) |
| SELECT_n_of_all_aggr_func | [0.00, 3.00, 14.00] 3.63 / 3.11 |
| FROM_n_of_join_paths | |
| 1 | 164 (16%) |
| 2 | 171 (17%) |
| 3 | 186 (19%) |
| 4 | 144 (14%) |
| 5 | 159 (16%) |
| 6 | 176 (18%) |
| FROM_n_of_super_joins__RIGHT | |
| 0 | 544 (54%) |
| 1 | 314 (31%) |
| 2 | 109 (11%) |
| 3 | 31 (3.1%) |
| 4 | 1 (0.1%) |
| 5 | 1 (0.1%) |
| FROM_n_of_super_joins__LEFT | |
| 0 | 541 (54%) |
| 1 | 314 (31%) |
| 2 | 116 (12%) |
| 3 | 25 (2.5%) |
| 4 | 3 (0.3%) |
| 5 | 1 (0.1%) |
| FROM_n_of_super_joins__FULL | |
| 0 | 554 (55%) |
| 1 | 317 (32%) |
| 2 | 105 (11%) |
| 3 | 21 (2.1%) |
| 4 | 3 (0.3%) |
| FROM_n_of_joins__INNER | [0.00, 3.00, 13.00] 3.28 / 2.49 |
| FROM_n_of_joins__RIGHT | [0.00, 3.00, 13.00] 3.27 / 2.46 |
| FROM_n_of_processed_rows | [5, 93,258, 433,730] 119,327 / 97,267 |
| WHERE_n_of_predicates | [0.00, 4.00, 12.00] 4.00 / 2.97 |
| WHERE_n_of_attrbs_of_type__character_varying | |

| Query parameter or value | Descriptive statistics |
|---------------------------------------|---------------------------------|
| 0 | 352 (35%) |
| 1 | 291 (29%) |
| 2 | 187 (19%) |
| 3 | 99 (9.9%) |
| 4 | 48 (4.8%) |
| 5 | 20 (2.0%) |
| 6 | 3 (0.3%) |
| WHERE_n_of_attribs_of_type__date | |
| 0 | 774 (77%) |
| 1 | 181 (18%) |
| 2 | 36 (3.6%) |
| 3 | 8 (0.8%) |
| 4 | 1 (0.1%) |
| WHERE_n_of_attribs_of_type__integer | |
| 0 | 349 (35%) |
| 1 | 289 (29%) |
| 2 | 175 (18%) |
| 3 | 113 (11%) |
| 4 | 52 (5.2%) |
| 5 | 18 (1.8%) |
| 6 | 4 (0.4%) |
| WHERE_n_of_attribs_of_type__character | |
| 0 | 642 (64%) |
| 1 | 258 (26%) |
| 2 | 82 (8.2%) |
| 3 | 14 (1.4%) |
| 4 | 4 (0.4%) |
| WHERE_n_of_attribs_of_type__numeric | |
| 0 | 548 (55%) |
| 1 | 292 (29%) |
| 2 | 116 (12%) |
| 3 | 31 (3.1%) |
| 4 | 13 (1.3%) |
| WHERE_n_of_pkey_attribs | |
| 0 | 527 (53%) |
| 1 | 305 (31%) |
| 2 | 123 (12%) |
| 3 | 33 (3.3%) |
| 4 | 10 (1.0%) |
| 5 | 2 (0.2%) |
| WHERE_n_of_connect_OR | [0.00, 2.00, 10.00] 2.32 / 2.43 |
| WHERE_n_of_operators__between | |
| 0 | 504 (50%) |
| 1 | 310 (31%) |
| 2 | 127 (13%) |
| 3 | 40 (4.0%) |

| Query parameter or value | Descriptive statistics |
|--|------------------------|
| 4 | 16 (1.6%) |
| 5 | 3 (0.3%) |
| WHERE_n_of_operators___greater_or_less | |
| 0 | 312 (31%) |
| 1 | 263 (26%) |
| 2 | 196 (20%) |
| 3 | 122 (12%) |
| 4 | 60 (6.0%) |
| 5 | 28 (2.8%) |
| 6 | 16 (1.6%) |
| 7 | 1 (0.1%) |
| 8 | 2 (0.2%) |
| WHERE_n_of_operators___in | |
| 0 | 531 (53%) |
| 1 | 289 (29%) |
| 2 | 111 (11%) |
| 3 | 56 (5.6%) |
| 4 | 11 (1.1%) |
| 5 | 1 (0.1%) |
| 6 | 1 (0.1%) |
| WHERE_n_of_operators___like | |
| 0 | 794 (79%) |
| 1 | 175 (18%) |
| 2 | 28 (2.8%) |
| 3 | 3 (0.3%) |
| WHERE_n_of_non_aggr_func___LOG | |
| 0 | 926 (93%) |
| 1 | 67 (6.7%) |
| 2 | 7 (0.7%) |
| WHERE_n_of_non_aggr_func___FLOOR | |
| 0 | 961 (96%) |
| 1 | 38 (3.8%) |
| 2 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___MONTH | |
| 0 | 971 (97%) |
| 1 | 27 (2.7%) |
| 2 | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___SQRT | |
| 0 | 883 (88%) |
| 1 | 106 (11%) |
| 2 | 10 (1.0%) |
| 3 | 1 (0.1%) |
| WHERE_n_of_non_aggr_func___ROUND | |
| 0 | 956 (96%) |
| 1 | 44 (4.4%) |
| WHERE_n_of_non_aggr_func___TRUNC | |

| Query parameter or value | Descriptive statistics |
|---|--|
| 0 | 966 (97%) |
| 1 | 34 (3.4%) |
| WHERE_n_of_non_aggr_func___DOW | |
| 0 | 957 (96%) |
| 1 | 43 (4.3%) |
| WHERE_n_of_non_aggr_func___YEAR | |
| 0 | 962 (96%) |
| 1 | 38 (3.8%) |
| WHERE_n_of_non_aggr_func___ABS | |
| 0 | 954 (95%) |
| 1 | 44 (4.4%) |
| 2 | 2 (0.2%) |
| WHERE_n_of_non_aggr_func___DAY | |
| 0 | 965 (97%) |
| 1 | 34 (3.4%) |
| 2 | 1 (0.1%) |
| WHERE_n_of_all_non_aggr_func | |
| 0 | 643 (64%) |
| 1 | 226 (23%) |
| 2 | 101 (10%) |
| 3 | 24 (2.4%) |
| 4 | 6 (0.6%) |
| GROUP_BY_n_of_columns | [0.00, 3.00, 12.00] 3.10 / 2.85 |
| HAVING_n_of_main_predicates | |
| 0 | 284 (28%) |
| 1 | 88 (8.8%) |
| 2 | 143 (14%) |
| 3 | 199 (20%) |
| 4 | 146 (15%) |
| 5 | 97 (9.7%) |
| 6 | 43 (4.3%) |
| HAVING_n_of_main_predicates___scalar_subquery | |
| 0 | 413 (41%) |
| 1 | 176 (18%) |
| 2 | 221 (22%) |
| 3 | 190 (19%) |
| HAVING_n_of_subqueries___scalar_subquery | [0.0, 1.0, 20.0] 2.5 / 3.6 |
| HAVING_n_of_processed_rows_by_subqueries | [0, 29,063, 4,659,325] 299,104 / 606,998 |
| ORDER_BY_n_of_columns | |
| 0 | 292 (29%) |
| 1 | 344 (34%) |
| 2 | 222 (22%) |
| 3 | 100 (10%) |
| 4 | 20 (2.0%) |
| 5 | 9 (0.9%) |
| 6 | 9 (0.9%) |

| Query parameter or value | Descriptive statistics |
|--------------------------|---------------------------|
| γ | 4 (0.4%) |
| limit | [1, 471, 1,000] 477 / 294 |
| offset | [0, 0, 999] 231 / 313 |