

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

## Scenario: 1-1jp\_\_ng

## Scale factor: 5 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, 44] 10 / 9      |
| SELECT_n_of_non_aggr_func__DOW   |                        |
| 0                                | 885 (89%)              |
| 1                                | 87 (8.7%)              |
| 2                                | 23 (2.3%)              |
| 3                                | 5 (0.5%)               |
| SELECT_n_of_non_aggr_func__LOWER |                        |
| 0                                | 780 (78%)              |
| 1                                | 172 (17%)              |
| 2                                | 37 (3.7%)              |
| 3                                | 10 (1.0%)              |
| 4                                | 1 (0.1%)               |
| SELECT_n_of_non_aggr_func__LTRIM |                        |
| 0                                | 795 (80%)              |
| 1                                | 157 (16%)              |

| Query parameter or value          | Descriptive statistics |
|-----------------------------------|------------------------|
| 2                                 | 32 (3.2%)              |
| 3                                 | 14 (1.4%)              |
| 4                                 | 1 (0.1%)               |
| 5                                 | 1 (0.1%)               |
| SELECT_n_of_non_aggr_func__MONTH  |                        |
| 0                                 | 833 (83%)              |
| 1                                 | 123 (12%)              |
| 2                                 | 32 (3.2%)              |
| 3                                 | 12 (1.2%)              |
| SELECT_n_of_non_aggr_func__ROUND  |                        |
| 0                                 | 936 (94%)              |
| 1                                 | 62 (6.2%)              |
| 2                                 | 1 (0.1%)               |
| 3                                 | 1 (0.1%)               |
| SELECT_n_of_non_aggr_func__RTRIM  |                        |
| 0                                 | 787 (79%)              |
| 1                                 | 154 (15%)              |
| 2                                 | 47 (4.7%)              |
| 3                                 | 11 (1.1%)              |
| 5                                 | 1 (0.1%)               |
| SELECT_n_of_non_aggr_func__SQRT   |                        |
| 0                                 | 820 (82%)              |
| 1                                 | 144 (14%)              |
| 2                                 | 30 (3.0%)              |
| 3                                 | 5 (0.5%)               |
| 4                                 | 1 (0.1%)               |
| SELECT_n_of_non_aggr_func__YEAR   |                        |
| 0                                 | 854 (85%)              |
| 1                                 | 107 (11%)              |
| 2                                 | 33 (3.3%)              |
| 3                                 | 5 (0.5%)               |
| 4                                 | 1 (0.1%)               |
| SELECT_n_of_non_aggr_func__UPPER  |                        |
| 0                                 | 777 (78%)              |
| 1                                 | 166 (17%)              |
| 2                                 | 42 (4.2%)              |
| 3                                 | 15 (1.5%)              |
| SELECT_n_of_non_aggr_func__SUBSTR |                        |
| 0                                 | 768 (77%)              |
| 1                                 | 181 (18%)              |
| 2                                 | 39 (3.9%)              |
| 3                                 | 12 (1.2%)              |
| SELECT_n_of_non_aggr_func__TRUNC  |                        |
| 0                                 | 940 (94%)              |
| 1                                 | 55 (5.5%)              |
| 2                                 | 4 (0.4%)               |

| Query parameter or value         | Descriptive statistics     |
|----------------------------------|----------------------------|
| 3                                | 1 (0.1%)                   |
| SELECT_n_of_non_aggr_func__LOG   |                            |
| 0                                | 883 (88%)                  |
| 1                                | 103 (10%)                  |
| 2                                | 14 (1.4%)                  |
| SELECT_n_of_non_aggr_func__ABS   |                            |
| 0                                | 932 (93%)                  |
| 1                                | 64 (6.4%)                  |
| 2                                | 4 (0.4%)                   |
| SELECT_n_of_non_aggr_func__DAY   |                            |
| 0                                | 909 (91%)                  |
| 1                                | 76 (7.6%)                  |
| 2                                | 15 (1.5%)                  |
| SELECT_n_of_non_aggr_func__FLOOR |                            |
| 0                                | 946 (95%)                  |
| 1                                | 50 (5.0%)                  |
| 2                                | 4 (0.4%)                   |
| SELECT_n_of_all_non_aggr_func    | [0.0, 1.0, 21.0] 2.7 / 3.7 |
| SELECT_n_of_aggr_func__MAX       |                            |
| 0                                | 938 (94%)                  |
| 1                                | 20 (2.0%)                  |
| 2                                | 25 (2.5%)                  |
| 3                                | 10 (1.0%)                  |
| 4                                | 4 (0.4%)                   |
| 5                                | 1 (0.1%)                   |
| 6                                | 2 (0.2%)                   |
| SELECT_n_of_aggr_func__COUNT     |                            |
| 0                                | 942 (94%)                  |
| 1                                | 34 (3.4%)                  |
| 2                                | 15 (1.5%)                  |
| 3                                | 6 (0.6%)                   |
| 4                                | 1 (0.1%)                   |
| 5                                | 1 (0.1%)                   |
| 6                                | 1 (0.1%)                   |
| SELECT_n_of_aggr_func__MIN       |                            |
| 0                                | 932 (93%)                  |
| 1                                | 33 (3.3%)                  |
| 2                                | 17 (1.7%)                  |
| 3                                | 9 (0.9%)                   |
| 4                                | 6 (0.6%)                   |
| 5                                | 2 (0.2%)                   |
| 6                                | 1 (0.1%)                   |
| SELECT_n_of_aggr_func__SUM       |                            |
| 0                                | 982 (98%)                  |
| 1                                | 17 (1.7%)                  |
| 2                                | 1 (0.1%)                   |

| Query parameter or value                        | Descriptive statistics                             |
|---|--|
| SELECT__n_of__aggr_func__AVG                    |  |
| 0   | 989 (99%)  |
| 1   | 10 (1.0%)  |
| 2   | 1 (0.1%)   |
| SELECT__n_of__aggr_func__COUNT_DISTINCT         |  |
| 0   | 921 (92%)  |
| 1   | 36 (3.6%)  |
| 2   | 21 (2.1%)  |
| 3   | 13 (1.3%)  |
| 4   | 6 (0.6%)   |
| 5   | 2 (0.2%)   |
| 6   | 1 (0.1%)   |
| SELECT__n_of__all__aggr_func                    |  |
| 0   | 764 (76%)  |
| 1   | 79 (7.9%)  |
| 2   | 64 (6.4%)  |
| 3   | 53 (5.3%)  |
| 4   | 24 (2.4%)  |
| 5   | 9 (0.9%)   |
| 6   | 4 (0.4%)   |
| 7   | 3 (0.3%)   |
| FROM__n_of__join_paths                          |  |
| 1   | 1,000 (100%)                                       |
| FROM__n_of__joins__INNER                        |  |
| 0   | 490 (49%)  |
| 1   | 320 (32%)  |
| 2   | 138 (14%)  |
| 3   | 46 (4.6%)  |
| 4   | 6 (0.6%)   |
| FROM__n_of__joins__RIGHT                        |  |
| 0   | 498 (50%)  |
| 1   | 344 (34%)  |
| 2   | 128 (13%)  |
| 3   | 26 (2.6%)  |
| 4   | 4 (0.4%)   |
| FROM__n_of__processed_rows                      | [5, 5,000,323, 38,249,944] 13,780,793 / 15,396,357 |
| WHERE__n_of__predicates                         |  |
| 0   | 221 (22%)  |
| 1   | 246 (25%)  |
| 2   | 239 (24%)  |
| 3   | 160 (16%)  |
| 4   | 89 (8.9%)  |
| 5   | 34 (3.4%)  |
| 6   | 11 (1.1%)  |
| WHERE__n_of__attribs_of_type__character_varying |  |
| 0   | 549 (55%)  |

| Query parameter or value              | Descriptive statistics |
|---------------------------------------|------------------------|
| 1                                     | 305 (31%)              |
| 2                                     | 124 (12%)              |
| 3                                     | 22 (2.2%)              |
| WHERE_n_of_attribs_of_type__character |                        |
| 0                                     | 801 (80%)              |
| 1                                     | 173 (17%)              |
| 2                                     | 24 (2.4%)              |
| 3                                     | 2 (0.2%)               |
| WHERE_n_of_attribs_of_type__integer   |                        |
| 0                                     | 559 (56%)              |
| 1                                     | 340 (34%)              |
| 2                                     | 87 (8.7%)              |
| 3                                     | 13 (1.3%)              |
| 4                                     | 1 (0.1%)               |
| WHERE_n_of_attribs_of_type__date      |                        |
| 0                                     | 903 (90%)              |
| 1                                     | 90 (9.0%)              |
| 2                                     | 7 (0.7%)               |
| WHERE_n_of_attribs_of_type__numeric   |                        |
| 0                                     | 751 (75%)              |
| 1                                     | 213 (21%)              |
| 2                                     | 32 (3.2%)              |
| 3                                     | 4 (0.4%)               |
| WHERE_n_of_pkey_attribs               |                        |
| 0                                     | 725 (73%)              |
| 1                                     | 242 (24%)              |
| 2                                     | 31 (3.1%)              |
| 3                                     | 2 (0.2%)               |
| WHERE_n_of_connect_OR                 |                        |
| 0                                     | 625 (63%)              |
| 1                                     | 210 (21%)              |
| 2                                     | 108 (11%)              |
| 3                                     | 37 (3.7%)              |
| 4                                     | 18 (1.8%)              |
| 5                                     | 2 (0.2%)               |
| WHERE_n_of_operators__greater_or_less |                        |
| 0                                     | 505 (51%)              |
| 1                                     | 350 (35%)              |
| 2                                     | 104 (10%)              |
| 3                                     | 31 (3.1%)              |
| 4                                     | 8 (0.8%)               |
| 5                                     | 2 (0.2%)               |
| WHERE_n_of_operators__like            |                        |
| 0                                     | 878 (88%)              |
| 1                                     | 116 (12%)              |
| 2                                     | 6 (0.6%)               |

| Query parameter or value        | Descriptive statistics |
|---------------------------------|------------------------|
| WHERE_n_of_operators__between   |                        |
| 0                               | 728 (73%)              |
| 1                               | 222 (22%)              |
| 2                               | 46 (4.6%)              |
| 3                               | 2 (0.2%)               |
| 4                               | 2 (0.2%)               |
| WHERE_n_of_operators__in        |                        |
| 0                               | 729 (73%)              |
| 1                               | 213 (21%)              |
| 2                               | 54 (5.4%)              |
| 3                               | 4 (0.4%)               |
| WHERE_n_of_non_aggr_func__LOG   |                        |
| 0                               | 963 (96%)              |
| 1                               | 35 (3.5%)              |
| 2                               | 2 (0.2%)               |
| WHERE_n_of_non_aggr_func__SQRT  |                        |
| 0                               | 944 (94%)              |
| 1                               | 54 (5.4%)              |
| 2                               | 2 (0.2%)               |
| WHERE_n_of_non_aggr_func__TRUNC |                        |
| 0                               | 977 (98%)              |
| 1                               | 23 (2.3%)              |
| WHERE_n_of_non_aggr_func__ROUND |                        |
| 0                               | 985 (99%)              |
| 1                               | 15 (1.5%)              |
| WHERE_n_of_non_aggr_func__YEAR  |                        |
| 0                               | 986 (99%)              |
| 1                               | 14 (1.4%)              |
| WHERE_n_of_non_aggr_func__ABS   |                        |
| 0                               | 984 (98%)              |
| 1                               | 15 (1.5%)              |
| 2                               | 1 (0.1%)               |
| WHERE_n_of_non_aggr_func__DAY   |                        |
| 0                               | 984 (98%)              |
| 1                               | 16 (1.6%)              |
| WHERE_n_of_non_aggr_func__MONTH |                        |
| 0                               | 985 (99%)              |
| 1                               | 15 (1.5%)              |
| WHERE_n_of_non_aggr_func__FLOOR |                        |
| 0                               | 984 (98%)              |
| 1                               | 16 (1.6%)              |
| WHERE_n_of_non_aggr_func__DOW   |                        |
| 0                               | 991 (99%)              |
| 1                               | 9 (0.9%)               |
| WHERE_n_of_all_non_aggr_func    |                        |
| 0                               | 810 (81%)              |

| Query parameter or value    | Descriptive statistics    |
|-----------------------------|---------------------------|
| <i>1</i>                    | 161 (16%)                 |
| <i>2</i>                    | 26 (2.6%)                 |
| <i>3</i>                    | 3 (0.3%)                  |
| GROUP_BY_n_of_columns       |                           |
| <i>0</i>                    | 1,000 (100%)              |
| HAVING_n_of_main_predicates |                           |
| <i>0</i>                    | 1,000 (100%)              |
| ORDER_BY_n_of_columns       |                           |
| <i>0</i>                    | 430 (43%)                 |
| <i>1</i>                    | 298 (30%)                 |
| <i>2</i>                    | 170 (17%)                 |
| <i>3</i>                    | 66 (6.6%)                 |
| <i>4</i>                    | 30 (3.0%)                 |
| <i>5</i>                    | 5 (0.5%)                  |
| <i>6</i>                    | 1 (0.1%)                  |
| limit                       | [1, 493, 1,000] 494 / 291 |
| offset                      | [0, 0, 999] 255 / 329     |