

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

```
## Scenario: 1-6jp_ujpp__g__h0-3ss_h0nss

## Scale factor: 100 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, 8, 173] 16 / 25    |
| SELECT_n_of_non_aggr_func__DAY |                        |
| 0                              | 872 (87%)              |
| 1                              | 91 (9.1%)              |
| 2                              | 20 (2.0%)              |
| 3                              | 10 (1.0%)              |
| 4                              | 5 (0.5%)               |
| 5                              | 2 (0.2%)               |
| SELECT_n_of_non_aggr_func__DOW |                        |
| 0                              | 860 (86%)              |
| 1                              | 94 (9.4%)              |
| 2                              | 30 (3.0%)              |
| 3                              | 9 (0.9%)               |
| 4                              | 3 (0.3%)               |
| 5                              | 2 (0.2%)               |

| Query parameter or value          | Descriptive statistics          |
|-----------------------------------|---------------------------------|
| 6                                 | 2 (0.2%)                        |
| SELECT_n_of_non_aggr_func__FLOOR  |                                 |
| 0                                 | 909 (91%)                       |
| 1                                 | 74 (7.4%)                       |
| 2                                 | 14 (1.4%)                       |
| 3                                 | 2 (0.2%)                        |
| 4                                 | 1 (0.1%)                        |
| SELECT_n_of_non_aggr_func__LOG    |                                 |
| 0                                 | 849 (85%)                       |
| 1                                 | 116 (12%)                       |
| 2                                 | 21 (2.1%)                       |
| 3                                 | 7 (0.7%)                        |
| 4                                 | 5 (0.5%)                        |
| 5                                 | 2 (0.2%)                        |
| SELECT_n_of_non_aggr_func__MONTH  |                                 |
| 0                                 | 798 (80%)                       |
| 1                                 | 137 (14%)                       |
| 2                                 | 35 (3.5%)                       |
| 3                                 | 11 (1.1%)                       |
| 4                                 | 5 (0.5%)                        |
| 5                                 | 8 (0.8%)                        |
| 6                                 | 3 (0.3%)                        |
| 7                                 | 1 (0.1%)                        |
| 8                                 | 2 (0.2%)                        |
| SELECT_n_of_non_aggr_func__LOWER  | [0.00, 0.00, 9.00] 0.58 / 1.02  |
| SELECT_n_of_non_aggr_func__SUBSTR | [0.00, 0.00, 10.00] 0.60 / 1.08 |
| SELECT_n_of_non_aggr_func__TRUNC  |                                 |
| 0                                 | 921 (92%)                       |
| 1                                 | 62 (6.2%)                       |
| 2                                 | 12 (1.2%)                       |
| 3                                 | 4 (0.4%)                        |
| 4                                 | 1 (0.1%)                        |
| SELECT_n_of_non_aggr_func__SQRT   | [0.00, 0.00, 10.00] 0.35 / 0.99 |
| SELECT_n_of_non_aggr_func__LTRIM  | [0.00, 0.00, 9.00] 0.62 / 1.11  |
| SELECT_n_of_non_aggr_func__YEAR   |                                 |
| 0                                 | 833 (83%)                       |
| 1                                 | 110 (11%)                       |
| 2                                 | 27 (2.7%)                       |
| 3                                 | 13 (1.3%)                       |
| 4                                 | 7 (0.7%)                        |
| 5                                 | 4 (0.4%)                        |
| 6                                 | 5 (0.5%)                        |
| 7                                 | 1 (0.1%)                        |
| SELECT_n_of_non_aggr_func__ROUND  |                                 |
| 0                                 | 927 (93%)                       |
| 1                                 | 60 (6.0%)                       |

| Query parameter or value              | Descriptive statistics          |
|---------------------------------------|---------------------------------|
| 2                                     | 8 (0.8%)                        |
| 3                                     | 3 (0.3%)                        |
| 4                                     | 1 (0.1%)                        |
| 5                                     | 1 (0.1%)                        |
| SELECT_n_of_non_aggr_func__RTRIM      | [0.00, 0.00, 11.00] 0.64 / 1.14 |
| SELECT_n_of_non_aggr_func__UPPER      | [0.00, 0.00, 9.00] 0.61 / 1.09  |
| SELECT_n_of_non_aggr_func__ABS        |                                 |
| 0                                     | 920 (92%)                       |
| 1                                     | 65 (6.5%)                       |
| 2                                     | 14 (1.4%)                       |
| 4                                     | 1 (0.1%)                        |
| SELECT_n_of_all_non_aggr_func         | [0.0, 3.0, 74.0] 5.0 / 8.4      |
| SELECT_n_of_aggr_func__COUNT_DISTINCT | [0.00, 0.00, 12.00] 0.88 / 2.03 |
| SELECT_n_of_aggr_func__COUNT          | [0.00, 0.00, 14.00] 0.85 / 1.94 |
| SELECT_n_of_aggr_func__SUM            |                                 |
| 0                                     | 926 (93%)                       |
| 1                                     | 50 (5.0%)                       |
| 2                                     | 19 (1.9%)                       |
| 3                                     | 4 (0.4%)                        |
| 5                                     | 1 (0.1%)                        |
| SELECT_n_of_aggr_func__MAX            | [0.00, 0.00, 14.00] 0.90 / 2.03 |
| SELECT_n_of_aggr_func__MIN            | [0.00, 0.00, 14.00] 0.78 / 1.88 |
| SELECT_n_of_aggr_func__AVG            |                                 |
| 0                                     | 941 (94%)                       |
| 1                                     | 39 (3.9%)                       |
| 2                                     | 13 (1.3%)                       |
| 3                                     | 6 (0.6%)                        |
| 4                                     | 1 (0.1%)                        |
| SELECT_n_of_all_aggr_func             | [0.0, 3.0, 14.0] 3.6 / 3.2      |
| FROM_n_of_join_paths                  |                                 |
| 1                                     | 179 (18%)                       |
| 2                                     | 175 (18%)                       |
| 3                                     | 156 (16%)                       |
| 4                                     | 177 (18%)                       |
| 5                                     | 160 (16%)                       |
| 6                                     | 153 (15%)                       |
| FROM_n_of_super_joins__FULL           |                                 |
| 0                                     | 552 (55%)                       |
| 1                                     | 304 (30%)                       |
| 2                                     | 115 (12%)                       |
| 3                                     | 26 (2.6%)                       |
| 4                                     | 3 (0.3%)                        |
| FROM_n_of_super_joins__LEFT           |                                 |
| 0                                     | 557 (56%)                       |
| 1                                     | 311 (31%)                       |
| 2                                     | 106 (11%)                       |

| Query parameter or value                      | Descriptive statistics                                      |
|---|---|
| 3   | 26 (2.6%)   |
| FROM_n_of_super_joins__RIGHT                  |   |
| 0   | 571 (57%)   |
| 1   | 300 (30%)   |
| 2   | 107 (11%)   |
| 3   | 18 (1.8%)   |
| 4   | 4 (0.4%)  |
| FROM_n_of_joins__INNER                        | [0.00, 3.00, 12.00] 3.28 / 2.48                             |
| FROM_n_of_joins__RIGHT                        | [0.00, 3.00, 13.00] 3.22 / 2.47                             |
| FROM_n_of_processed_rows                      | [5, 928,538,025, 4,495,227,462] 1,173,165,546 / 948,381,914 |
| WHERE_n_of_predicates                         | [0.00, 3.00, 13.00] 3.74 / 2.85                             |
| WHERE_n_of_attribs_of_type__character         |   |
| 0   | 633 (63%)   |
| 1   | 273 (27%)   |
| 2   | 77 (7.7%)   |
| 3   | 16 (1.6%)   |
| 4   | 1 (0.1%)  |
| WHERE_n_of_attribs_of_type__character_varying |   |
| 0   | 382 (38%)   |
| 1   | 289 (29%)   |
| 2   | 187 (19%)   |
| 3   | 79 (7.9%)   |
| 4   | 39 (3.9%)   |
| 5   | 18 (1.8%)   |
| 6   | 4 (0.4%)  |
| 7   | 2 (0.2%)  |
| WHERE_n_of_attribs_of_type__integer           |   |
| 0   | 370 (37%)   |
| 1   | 291 (29%)   |
| 2   | 185 (19%)   |
| 3   | 93 (9.3%)   |
| 4   | 43 (4.3%)   |
| 5   | 13 (1.3%)   |
| 6   | 5 (0.5%)  |
| WHERE_n_of_attribs_of_type__date              |   |
| 0   | 801 (80%)   |
| 1   | 158 (16%)   |
| 2   | 38 (3.8%)   |
| 3   | 2 (0.2%)  |
| 4   | 1 (0.1%)  |
| WHERE_n_of_attribs_of_type__numeric           |   |
| 0   | 568 (57%)   |
| 1   | 287 (29%)   |
| 2   | 114 (11%)   |
| 3   | 21 (2.1%)   |

| Query parameter or value              | Descriptive statistics          |
|---------------------------------------|---------------------------------|
| 4                                     | 7 (0.7%)                        |
| 5                                     | 3 (0.3%)                        |
| WHERE_n_of_pkey_attris                |                                 |
| 0                                     | 517 (52%)                       |
| 1                                     | 323 (32%)                       |
| 2                                     | 108 (11%)                       |
| 3                                     | 46 (4.6%)                       |
| 4                                     | 6 (0.6%)                        |
| WHERE_n_of_connect_OR                 | [0.00, 1.00, 11.00] 2.10 / 2.30 |
| WHERE_n_of_operators__between         |                                 |
| 0                                     | 526 (53%)                       |
| 1                                     | 313 (31%)                       |
| 2                                     | 107 (11%)                       |
| 3                                     | 37 (3.7%)                       |
| 4                                     | 12 (1.2%)                       |
| 5                                     | 5 (0.5%)                        |
| WHERE_n_of_operators__greater_or_less |                                 |
| 0                                     | 343 (34%)                       |
| 1                                     | 268 (27%)                       |
| 2                                     | 189 (19%)                       |
| 3                                     | 109 (11%)                       |
| 4                                     | 57 (5.7%)                       |
| 5                                     | 24 (2.4%)                       |
| 6                                     | 6 (0.6%)                        |
| 7                                     | 2 (0.2%)                        |
| 8                                     | 2 (0.2%)                        |
| WHERE_n_of_operators__in              |                                 |
| 0                                     | 555 (56%)                       |
| 1                                     | 275 (28%)                       |
| 2                                     | 121 (12%)                       |
| 3                                     | 34 (3.4%)                       |
| 4                                     | 12 (1.2%)                       |
| 5                                     | 2 (0.2%)                        |
| 7                                     | 1 (0.1%)                        |
| WHERE_n_of_operators__like            |                                 |
| 0                                     | 796 (80%)                       |
| 1                                     | 179 (18%)                       |
| 2                                     | 22 (2.2%)                       |
| 3                                     | 3 (0.3%)                        |
| WHERE_n_of_non_aggr_func__DOW         |                                 |
| 0                                     | 975 (98%)                       |
| 1                                     | 25 (2.5%)                       |
| WHERE_n_of_non_aggr_func__SQRT        |                                 |
| 0                                     | 881 (88%)                       |
| 1                                     | 117 (12%)                       |
| 2                                     | 2 (0.2%)                        |

| Query parameter or value                     | Descriptive statistics          |
|--|---------------------------------|
| WHERE_n_of_non_aggr_func__ABS                |                                 |
| 0  | 958 (96%)                       |
| 1  | 41 (4.1%)                       |
| 2  | 1 (0.1%)                        |
| WHERE_n_of_non_aggr_func__MONTH              |                                 |
| 0  | 967 (97%)                       |
| 1  | 33 (3.3%)                       |
| WHERE_n_of_non_aggr_func__LOG                |                                 |
| 0  | 916 (92%)                       |
| 1  | 79 (7.9%)                       |
| 2  | 5 (0.5%)                        |
| WHERE_n_of_non_aggr_func__ROUND              |                                 |
| 0  | 957 (96%)                       |
| 1  | 42 (4.2%)                       |
| 2  | 1 (0.1%)                        |
| WHERE_n_of_non_aggr_func__FLOOR              |                                 |
| 0  | 956 (96%)                       |
| 1  | 42 (4.2%)                       |
| 2  | 2 (0.2%)                        |
| WHERE_n_of_non_aggr_func__TRUNC              |                                 |
| 0  | 967 (97%)                       |
| 1  | 33 (3.3%)                       |
| WHERE_n_of_non_aggr_func__DAY                |                                 |
| 0  | 964 (96%)                       |
| 1  | 35 (3.5%)                       |
| 2  | 1 (0.1%)                        |
| WHERE_n_of_non_aggr_func__YEAR               |                                 |
| 0  | 963 (96%)                       |
| 1  | 34 (3.4%)                       |
| 2  | 3 (0.3%)                        |
| WHERE_n_of_all_non_aggr_func                 |                                 |
| 0  | 620 (62%)                       |
| 1  | 269 (27%)                       |
| 2  | 91 (9.1%)                       |
| 3  | 20 (2.0%)                       |
| GROUP_BY_n_of_columns                        | [0.00, 2.00, 12.00] 3.05 / 2.89 |
| HAVING_n_of_main_predicates                  |                                 |
| 0  | 300 (30%)                       |
| 1  | 103 (10%)                       |
| 2  | 142 (14%)                       |
| 3  | 191 (19%)                       |
| 4  | 126 (13%)                       |
| 5  | 95 (9.5%)                       |
| 6  | 43 (4.3%)                       |
| HAVING_n_of_main_predicates__scalar_subquery |                                 |
| 0  | 444 (44%)                       |

| Query parameter or value                 | Descriptive statistics   |
|--|--|
| <i>1</i>                                 | 183 (18%)  |
| <i>2</i>                                 | 198 (20%)  |
| <i>3</i>                                 | 175 (18%)  |
| HAVING_n_of_subqueries__scalar_subquery  | [0.0, 1.0, 18.0] 2.2 / 3.3                                     |
| HAVING_n_of_processed_rows_by_subqueries | [0, 141,000,120, 33,931,478,568] 2,348,403,918 / 4,708,674,732 |
| ORDER_BY_n_of_columns                    |  |
| <i>0</i>                                 | 309 (31%)  |
| <i>1</i>                                 | 316 (32%)  |
| <i>2</i>                                 | 244 (24%)  |
| <i>3</i>                                 | 90 (9.0%)  |
| <i>4</i>                                 | 28 (2.8%)  |
| <i>5</i>                                 | 9 (0.9%)   |
| <i>6</i>                                 | 2 (0.2%)   |
| <i>7</i>                                 | 2 (0.2%)   |
| limit                                    | [3, 526, 1,000] 521 / 290                                      |
| offset                                   | [0, 10, 998] 250 / 317   |