

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

## Scenario: 1-6jp\_rjpp\_\_g\_\_h0-3ss\_h0nss

## Scale factor: 1000 GB

## The 1000-query pack was generated on: 2023-11-07

## 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, 166] 16 / 24             |
| SELECT_n_of_non_aggr_func__TRUNC |                                 |
| 0                                | 890 (89%)                       |
| 1                                | 94 (9.4%)                       |
| 2                                | 12 (1.2%)                       |
| 3                                | 4 (0.4%)                        |
| SELECT_n_of_non_aggr_func__DAY   |                                 |
| 0                                | 885 (89%)                       |
| 1                                | 79 (7.9%)                       |
| 2                                | 20 (2.0%)                       |
| 3                                | 9 (0.9%)                        |
| 4                                | 3 (0.3%)                        |
| 5                                | 3 (0.3%)                        |
| 7                                | 1 (0.1%)                        |
| SELECT_n_of_non_aggr_func__LOWER | [0.00, 0.00, 12.00] 0.64 / 1.14 |

| Query parameter or value            | Descriptive statistics          |
|-------------------------------------|---------------------------------|
| SELECT__n_of__non_aggr_func__LTRIM  |                                 |
| 0                                   | 624 (62%)                       |
| 1                                   | 246 (25%)                       |
| 2                                   | 72 (7.2%)                       |
| 3                                   | 28 (2.8%)                       |
| 4                                   | 15 (1.5%)                       |
| 5                                   | 7 (0.7%)                        |
| 6                                   | 4 (0.4%)                        |
| 7                                   | 3 (0.3%)                        |
| 8                                   | 1 (0.1%)                        |
| SELECT__n_of__non_aggr_func__MONTH  | [0.00, 0.00, 10.00] 0.39 / 1.08 |
| SELECT__n_of__non_aggr_func__RTRIM  |                                 |
| 0                                   | 598 (60%)                       |
| 1                                   | 270 (27%)                       |
| 2                                   | 82 (8.2%)                       |
| 3                                   | 24 (2.4%)                       |
| 4                                   | 9 (0.9%)                        |
| 5                                   | 9 (0.9%)                        |
| 6                                   | 6 (0.6%)                        |
| 7                                   | 2 (0.2%)                        |
| SELECT__n_of__non_aggr_func__SUBSTR |                                 |
| 0                                   | 632 (63%)                       |
| 1                                   | 245 (25%)                       |
| 2                                   | 72 (7.2%)                       |
| 3                                   | 22 (2.2%)                       |
| 4                                   | 12 (1.2%)                       |
| 5                                   | 11 (1.1%)                       |
| 6                                   | 2 (0.2%)                        |
| 7                                   | 3 (0.3%)                        |
| 8                                   | 1 (0.1%)                        |
| SELECT__n_of__non_aggr_func__YEAR   |                                 |
| 0                                   | 825 (83%)                       |
| 1                                   | 113 (11%)                       |
| 2                                   | 31 (3.1%)                       |
| 3                                   | 15 (1.5%)                       |
| 4                                   | 10 (1.0%)                       |
| 5                                   | 1 (0.1%)                        |
| 6                                   | 2 (0.2%)                        |
| 7                                   | 2 (0.2%)                        |
| 8                                   | 1 (0.1%)                        |
| SELECT__n_of__non_aggr_func__DOW    |                                 |
| 0                                   | 829 (83%)                       |
| 1                                   | 120 (12%)                       |
| 2                                   | 26 (2.6%)                       |
| 3                                   | 15 (1.5%)                       |
| 4                                   | 6 (0.6%)                        |

| Query parameter or value               | Descriptive statistics          |
|--|---------------------------------|
| 5                                      | 2 (0.2%)                        |
| 6                                      | 2 (0.2%)                        |
| SELECT_n_of_non_aggr_func___FLOOR      |                                 |
| 0                                      | 925 (93%)                       |
| 1                                      | 58 (5.8%)                       |
| 2                                      | 14 (1.4%)                       |
| 3                                      | 3 (0.3%)                        |
| SELECT_n_of_non_aggr_func___LOG        |                                 |
| 0                                      | 857 (86%)                       |
| 1                                      | 115 (12%)                       |
| 2                                      | 13 (1.3%)                       |
| 3                                      | 6 (0.6%)                        |
| 4                                      | 6 (0.6%)                        |
| 5                                      | 1 (0.1%)                        |
| 7                                      | 1 (0.1%)                        |
| 8                                      | 1 (0.1%)                        |
| SELECT_n_of_non_aggr_func___ROUND      |                                 |
| 0                                      | 911 (91%)                       |
| 1                                      | 73 (7.3%)                       |
| 2                                      | 13 (1.3%)                       |
| 3                                      | 3 (0.3%)                        |
| SELECT_n_of_non_aggr_func___SQRT       |                                 |
| 0                                      | 785 (79%)                       |
| 1                                      | 141 (14%)                       |
| 2                                      | 38 (3.8%)                       |
| 3                                      | 15 (1.5%)                       |
| 4                                      | 9 (0.9%)                        |
| 5                                      | 4 (0.4%)                        |
| 6                                      | 4 (0.4%)                        |
| 7                                      | 2 (0.2%)                        |
| 8                                      | 2 (0.2%)                        |
| SELECT_n_of_non_aggr_func___UPPER      | [0.00, 0.00, 9.00] 0.67 / 1.16  |
| SELECT_n_of_non_aggr_func___ABS        |                                 |
| 0                                      | 917 (92%)                       |
| 1                                      | 63 (6.3%)                       |
| 2                                      | 14 (1.4%)                       |
| 3                                      | 6 (0.6%)                        |
| SELECT_n_of_all_non_aggr_func          | [0, 3, 63] 5 / 9                |
| SELECT_n_of_aggr_func___COUNT          | [0.00, 0.00, 12.00] 0.97 / 2.05 |
| SELECT_n_of_aggr_func___COUNT_DISTINCT | [0.00, 0.00, 13.00] 0.71 / 1.77 |
| SELECT_n_of_aggr_func___MAX            | [0.00, 0.00, 13.00] 0.83 / 1.92 |
| SELECT_n_of_aggr_func___AVG            |                                 |
| 0                                      | 927 (93%)                       |
| 1                                      | 52 (5.2%)                       |
| 2                                      | 16 (1.6%)                       |
| 3                                      | 3 (0.3%)                        |

| Query parameter or value                            | Descriptive statistics  |
|---|---|
| 4   | 1 (0.1%)  |
| 5   | 1 (0.1%)  |
| SELECT__n__of__aggr__func__MIN                      | [0.00, 0.00, 13.00] 0.84 / 1.91                                   |
| SELECT__n__of__aggr__func__SUM                      |   |
| 0   | 931 (93%)   |
| 1   | 44 (4.4%)   |
| 2   | 19 (1.9%)   |
| 3   | 4 (0.4%)  |
| 4   | 1 (0.1%)  |
| 5   | 1 (0.1%)  |
| SELECT__n__of__all__aggr__func                      | [0.00, 3.00, 14.00] 3.55 / 3.05                                   |
| FROM__n__of__join__paths                            |   |
| 1   | 174 (17%)   |
| 2   | 153 (15%)   |
| 3   | 179 (18%)   |
| 4   | 159 (16%)   |
| 5   | 158 (16%)   |
| 6   | 177 (18%)   |
| FROM__n__of__super__joins__LEFT                     |   |
| 0   | 562 (56%)   |
| 1   | 305 (31%)   |
| 2   | 114 (11%)   |
| 3   | 17 (1.7%)   |
| 4   | 2 (0.2%)  |
| FROM__n__of__super__joins__FULL                     |   |
| 0   | 536 (54%)   |
| 1   | 329 (33%)   |
| 2   | 110 (11%)   |
| 3   | 21 (2.1%)   |
| 4   | 4 (0.4%)  |
| FROM__n__of__super__joins__RIGHT                    |   |
| 0   | 541 (54%)   |
| 1   | 314 (31%)   |
| 2   | 122 (12%)   |
| 3   | 22 (2.2%)   |
| 4   | 1 (0.1%)  |
| FROM__n__of__joins__INNER                           | [0.00, 3.00, 14.00] 3.36 / 2.57                                   |
| FROM__n__of__joins__RIGHT                           | [0.00, 3.00, 13.00] 3.31 / 2.54                                   |
| FROM__n__of__processed__rows                        | [5, 9,309,989,797, 45,899,938,359] 12,096,074,602 / 9,947,518,267 |
| WHERE__n__of__predicates                            | [0.00, 4.00, 13.00] 3.95 / 3.01                                   |
| WHERE__n__of__attribs__of__type__character__varying |   |
| 0   | 357 (36%)   |
| 1   | 286 (29%)   |
| 2   | 208 (21%)   |
| 3   | 85 (8.5%)   |

| Query parameter or value              | Descriptive statistics          |
|---------------------------------------|---------------------------------|
| 4                                     | 41 (4.1%)                       |
| 5                                     | 18 (1.8%)                       |
| 6                                     | 3 (0.3%)                        |
| 7                                     | 1 (0.1%)                        |
| 8                                     | 1 (0.1%)                        |
| WHERE_n_of_attribs_of_type__character |                                 |
| 0                                     | 638 (64%)                       |
| 1                                     | 253 (25%)                       |
| 2                                     | 86 (8.6%)                       |
| 3                                     | 19 (1.9%)                       |
| 4                                     | 4 (0.4%)                        |
| WHERE_n_of_attribs_of_type__integer   |                                 |
| 0                                     | 372 (37%)                       |
| 1                                     | 281 (28%)                       |
| 2                                     | 169 (17%)                       |
| 3                                     | 108 (11%)                       |
| 4                                     | 46 (4.6%)                       |
| 5                                     | 17 (1.7%)                       |
| 6                                     | 6 (0.6%)                        |
| 7                                     | 1 (0.1%)                        |
| WHERE_n_of_attribs_of_type__numeric   |                                 |
| 0                                     | 546 (55%)                       |
| 1                                     | 297 (30%)                       |
| 2                                     | 101 (10%)                       |
| 3                                     | 45 (4.5%)                       |
| 4                                     | 9 (0.9%)                        |
| 5                                     | 1 (0.1%)                        |
| 6                                     | 1 (0.1%)                        |
| WHERE_n_of_attribs_of_type__date      |                                 |
| 0                                     | 772 (77%)                       |
| 1                                     | 190 (19%)                       |
| 2                                     | 31 (3.1%)                       |
| 3                                     | 6 (0.6%)                        |
| 4                                     | 1 (0.1%)                        |
| WHERE_n_of_pkey_attribs               |                                 |
| 0                                     | 537 (54%)                       |
| 1                                     | 288 (29%)                       |
| 2                                     | 111 (11%)                       |
| 3                                     | 49 (4.9%)                       |
| 4                                     | 8 (0.8%)                        |
| 5                                     | 7 (0.7%)                        |
| WHERE_n_of_connect_OR                 | [0.00, 2.00, 11.00] 2.29 / 2.45 |
| WHERE_n_of_operators__greater_or_less |                                 |
| 0                                     | 330 (33%)                       |
| 1                                     | 250 (25%)                       |
| 2                                     | 187 (19%)                       |

| Query parameter or value        | Descriptive statistics |
|---------------------------------|------------------------|
| 3                               | 124 (12%)              |
| 4                               | 69 (6.9%)              |
| 5                               | 27 (2.7%)              |
| 6                               | 12 (1.2%)              |
| 7                               | 1 (0.1%)               |
| WHERE_n_of_operators__between   |                        |
| 0                               | 510 (51%)              |
| 1                               | 314 (31%)              |
| 2                               | 110 (11%)              |
| 3                               | 49 (4.9%)              |
| 4                               | 15 (1.5%)              |
| 5                               | 2 (0.2%)               |
| WHERE_n_of_operators__in        |                        |
| 0                               | 537 (54%)              |
| 1                               | 290 (29%)              |
| 2                               | 115 (12%)              |
| 3                               | 44 (4.4%)              |
| 4                               | 12 (1.2%)              |
| 5                               | 2 (0.2%)               |
| WHERE_n_of_operators__like      |                        |
| 0                               | 775 (78%)              |
| 1                               | 196 (20%)              |
| 2                               | 26 (2.6%)              |
| 3                               | 3 (0.3%)               |
| WHERE_n_of_non_aggr_func__LOG   |                        |
| 0                               | 928 (93%)              |
| 1                               | 70 (7.0%)              |
| 2                               | 2 (0.2%)               |
| WHERE_n_of_non_aggr_func__ROUND |                        |
| 0                               | 959 (96%)              |
| 1                               | 40 (4.0%)              |
| 2                               | 1 (0.1%)               |
| WHERE_n_of_non_aggr_func__ABS   |                        |
| 0                               | 948 (95%)              |
| 1                               | 50 (5.0%)              |
| 2                               | 2 (0.2%)               |
| WHERE_n_of_non_aggr_func__DOW   |                        |
| 0                               | 964 (96%)              |
| 1                               | 35 (3.5%)              |
| 2                               | 1 (0.1%)               |
| WHERE_n_of_non_aggr_func__SQRT  |                        |
| 0                               | 865 (87%)              |
| 1                               | 123 (12%)              |
| 2                               | 12 (1.2%)              |
| WHERE_n_of_non_aggr_func__FLOOR |                        |
| 0                               | 957 (96%)              |

| Query parameter or value                     | Descriptive statistics   |
|--|--|
| 1  | 42 (4.2%)  |
| 2  | 1 (0.1%)   |
| WHERE_n_of_non_aggr_func__TRUNC              |  |
| 0  | 960 (96%)  |
| 1  | 39 (3.9%)  |
| 2  | 1 (0.1%)   |
| WHERE_n_of_non_aggr_func__MONTH              |  |
| 0  | 960 (96%)  |
| 1  | 39 (3.9%)  |
| 2  | 1 (0.1%)   |
| WHERE_n_of_non_aggr_func__YEAR               |  |
| 0  | 968 (97%)  |
| 1  | 32 (3.2%)  |
| WHERE_n_of_non_aggr_func__DAY                |  |
| 0  | 964 (96%)  |
| 1  | 36 (3.6%)  |
| WHERE_n_of_all_non_aggr_func                 |  |
| 0  | 611 (61%)  |
| 1  | 261 (26%)  |
| 2  | 103 (10%)  |
| 3  | 21 (2.1%)  |
| 4  | 2 (0.2%)   |
| 5  | 2 (0.2%)   |
| GROUP_BY_n_of_columns                        | [0.00, 2.00, 13.00] 2.94 / 2.81                                      |
| HAVING_n_of_main_predicates                  |  |
| 0  | 308 (31%)  |
| 1  | 101 (10%)  |
| 2  | 157 (16%)  |
| 3  | 174 (17%)  |
| 4  | 136 (14%)  |
| 5  | 92 (9.2%)  |
| 6  | 32 (3.2%)  |
| HAVING_n_of_main_predicates__scalar_subquery |  |
| 0  | 440 (44%)  |
| 1  | 199 (20%)  |
| 2  | 204 (20%)  |
| 3  | 157 (16%)  |
| HAVING_n_of_subqueries__scalar_subquery      | [0.0, 1.0, 26.0] 2.3 / 3.6   |
| HAVING_n_of_processed_rows_by_subquery       | [0, 1050,000,210, 1,176,498,397,464] 28,880,483,064 / 69,556,328,986 |
| ORDER_BY_n_of_columns                        |  |
| 0  | 332 (33%)  |
| 1  | 322 (32%)  |
| 2  | 233 (23%)  |
| 3  | 76 (7.6%)  |
| 4  | 20 (2.0%)  |

| Query parameter or value | Descriptive statistics  |
|--------------------------|-------------------------|
| <i>5</i>                 | 9 (0.9%)                |
| <i>6</i>                 | 7 (0.7%)                |
| <i>9</i>                 | 1 (0.1%)                |
| limit                    | [3, 530, 999] 514 / 289 |
| offset                   | [0, 0, 997] 244 / 321   |