

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

Scenario: 1-6jp_rjpp__g__h0-3ss_h0-3nss

Scale factor: 2000 GB

The 1000-query pack was generated on: 2023-11-13

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, 189] 16 / 24
SELECT_n_of_non_aggr_func__LOWER	
0	609 (61%)
1	256 (26%)
2	85 (8.5%)
3	28 (2.8%)
4	8 (0.8%)
5	8 (0.8%)
6	2 (0.2%)
7	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	610 (61%)
1	266 (27%)
11	1 (0.1%)

Query parameter or value	Descriptive statistics
2	75 (7.5%)
3	27 (2.7%)
4	9 (0.9%)
5	7 (0.7%)
6	4 (0.4%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__SUBSTR	[0.00, 0.00, 11.00] 0.60 / 1.07
SELECT__n_of__non__aggr__func__UPPER	[0.00, 0.00, 11.00] 0.62 / 1.10
SELECT__n_of__non__aggr__func__DOW	
0	859 (86%)
1	105 (11%)
2	21 (2.1%)
3	8 (0.8%)
4	3 (0.3%)
5	4 (0.4%)
SELECT__n_of__non__aggr__func__FLOOR	
0	910 (91%)
1	76 (7.6%)
2	10 (1.0%)
3	3 (0.3%)
5	1 (0.1%)
SELECT__n_of__non__aggr__func__MONTH	
0	809 (81%)
1	134 (13%)
2	26 (2.6%)
3	9 (0.9%)
4	10 (1.0%)
5	5 (0.5%)
6	4 (0.4%)
7	2 (0.2%)
8	1 (0.1%)
SELECT__n_of__non__aggr__func__ROUND	
0	913 (91%)
1	67 (6.7%)
2	13 (1.3%)
3	5 (0.5%)
4	1 (0.1%)
6	1 (0.1%)
SELECT__n_of__non__aggr__func__RTRIM	
0	618 (62%)
1	271 (27%)
2	61 (6.1%)
3	23 (2.3%)
4	14 (1.4%)
5	7 (0.7%)
6	2 (0.2%)

Query parameter or value	Descriptive statistics
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___SQRT	
0	799 (80%)
1	140 (14%)
2	30 (3.0%)
3	14 (1.4%)
4	5 (0.5%)
5	7 (0.7%)
6	3 (0.3%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___TRUNC	
0	913 (91%)
1	75 (7.5%)
2	10 (1.0%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func___YEAR	
0	823 (82%)
1	129 (13%)
2	27 (2.7%)
3	10 (1.0%)
4	6 (0.6%)
5	2 (0.2%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___DAY	
0	881 (88%)
1	89 (8.9%)
2	18 (1.8%)
3	10 (1.0%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	862 (86%)
1	103 (10%)
2	20 (2.0%)
3	6 (0.6%)
4	8 (0.8%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	918 (92%)
1	67 (6.7%)
2	11 (1.1%)

Query parameter or value	Descriptive statistics
3	3 (0.3%)
5	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 73.0] 5.0 / 7.9
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 1.07 / 2.25
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 16.00] 0.91 / 1.96
SELECT_n_of_aggr_func__SUM	
0	936 (94%)
1	41 (4.1%)
2	18 (1.8%)
3	4 (0.4%)
4	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 12.00] 0.75 / 1.84
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 16.00] 0.83 / 1.99
SELECT_n_of_aggr_func__AVG	
0	926 (93%)
1	49 (4.9%)
2	17 (1.7%)
3	4 (0.4%)
4	4 (0.4%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 17.0] 3.8 / 3.2
FROM_n_of_join_paths	
1	161 (16%)
2	159 (16%)
3	162 (16%)
4	180 (18%)
5	156 (16%)
6	182 (18%)
FROM_n_of_super_joins__FULL	
0	526 (53%)
1	316 (32%)
2	118 (12%)
3	35 (3.5%)
4	5 (0.5%)
FROM_n_of_super_joins__LEFT	
0	515 (52%)
1	348 (35%)
2	105 (11%)
3	29 (2.9%)
4	3 (0.3%)
FROM_n_of_super_joins__RIGHT	
0	542 (54%)
1	314 (31%)
2	106 (11%)
3	35 (3.5%)
4	2 (0.2%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
FROM_n_of_joins__INNER	[0.00, 3.00, 13.00] 3.38 / 2.51
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.41 / 2.52
FROM_n_of_processed_rows	[5, 18,540,027,896, 88,440,166,937] 23,953,761,500 / 19,965,344,416
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.76 / 2.87
WHERE_n_of_attribs_of_type__character	
0	652 (65%)
1	253 (25%)
2	77 (7.7%)
3	18 (1.8%)
WHERE_n_of_attribs_of_type__character_varying	
0	384 (38%)
1	288 (29%)
2	185 (19%)
3	90 (9.0%)
4	37 (3.7%)
5	12 (1.2%)
6	4 (0.4%)
WHERE_n_of_attribs_of_type__integer	
0	369 (37%)
1	297 (30%)
2	179 (18%)
3	93 (9.3%)
4	42 (4.2%)
5	16 (1.6%)
6	3 (0.3%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type__date	
0	775 (78%)
1	184 (18%)
2	34 (3.4%)
3	7 (0.7%)
WHERE_n_of_attribs_of_type__numeric	
0	562 (56%)
1	283 (28%)
2	112 (11%)
3	29 (2.9%)
4	7 (0.7%)
5	7 (0.7%)
WHERE_n_of_pkey_attribs	
0	536 (54%)
1	302 (30%)
2	110 (11%)
3	39 (3.9%)
4	12 (1.2%)
5	1 (0.1%)

Query parameter or value	Descriptive statistics
WHERE_n_of_connect_OR	[0.00, 1.00, 11.00] 2.12 / 2.33
WHERE_n_of_operators__between	
0	528 (53%)
1	284 (28%)
2	124 (12%)
3	45 (4.5%)
4	11 (1.1%)
5	5 (0.5%)
6	3 (0.3%)
WHERE_n_of_operators__greater_or_less	
0	349 (35%)
1	269 (27%)
2	196 (20%)
3	87 (8.7%)
4	62 (6.2%)
5	26 (2.6%)
6	10 (1.0%)
8	1 (0.1%)
WHERE_n_of_operators__in	
0	537 (54%)
1	286 (29%)
2	125 (13%)
3	35 (3.5%)
4	16 (1.6%)
5	1 (0.1%)
WHERE_n_of_operators__like	
0	788 (79%)
1	179 (18%)
2	29 (2.9%)
3	4 (0.4%)
WHERE_n_of_non_aggr_func__SQRT	
0	862 (86%)
1	125 (13%)
2	13 (1.3%)
WHERE_n_of_non_aggr_func__MONTH	
0	961 (96%)
1	39 (3.9%)
WHERE_n_of_non_aggr_func__ABS	
0	957 (96%)
1	41 (4.1%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__YEAR	
0	959 (96%)
1	40 (4.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__FLOOR	

Query parameter or value	Descriptive statistics
0	960 (96%)
1	38 (3.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func__LOG	
0	923 (92%)
1	74 (7.4%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func__TRUNC	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__ROUND	
0	951 (95%)
1	45 (4.5%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func__DAY	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DOW	
0	955 (96%)
1	45 (4.5%)
WHERE_n_of_all_non_aggr_func	
0	592 (59%)
1	279 (28%)
2	92 (9.2%)
3	23 (2.3%)
4	13 (1.3%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 3.00, 14.00] 3.21 / 2.97
HAVING_n_of_main_predicates	[0.00, 4.00, 9.00] 3.45 / 2.58
HAVING_n_of_main_predicates__scalar_subquery	
0	426 (43%)
1	183 (18%)
2	204 (20%)
3	187 (19%)
HAVING_n_of_main_predicates__non_scalar_subquery	
0	438 (44%)
1	187 (19%)
2	194 (19%)
3	181 (18%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 22.0] 2.3 / 3.4
HAVING_n_of_subqueries__non_scalar_subquery	
0	438 (44%)
1	187 (19%)
2	194 (19%)

Query parameter or value	Descriptive statistics
<i>3</i>	181 (18%)
HAVING_n_of_processed_rows_by_subquery	[0, 27, 940, 055, 694, 871, 341, 586, 310] 82,881,904,207 / 131,483,032,133
ORDER_BY_n_of_columns	
<i>0</i>	301 (30%)
<i>1</i>	334 (33%)
<i>2</i>	227 (23%)
<i>3</i>	101 (10%)
<i>4</i>	21 (2.1%)
<i>5</i>	6 (0.6%)
<i>6</i>	7 (0.7%)
<i>7</i>	2 (0.2%)
<i>9</i>	1 (0.1%)
limit	[2, 501, 1,000] 509 / 293
offset	[0, 29, 995] 262 / 330