

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

Scenario: 1-6jp_ujpp__g__h0-3ss_h0nss

Scale factor: 2 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, 159] 16 / 24
SELECT_n_of_non_aggr_func___ABS	
0	904 (90%)
1	77 (7.7%)
2	14 (1.4%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___LOWER	
0	639 (64%)
1	238 (24%)
13	1 (0.1%)
2	76 (7.6%)
3	21 (2.1%)
4	12 (1.2%)
5	9 (0.9%)

Query parameter or value	Descriptive statistics
7	3 (0.3%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 10.00] 0.34 / 0.98
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 11.00] 0.59 / 1.08
SELECT_n_of_non_aggr_func__SQRT	
0	804 (80%)
1	132 (13%)
2	31 (3.1%)
3	13 (1.3%)
4	7 (0.7%)
5	8 (0.8%)
6	3 (0.3%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__SUBSTR	[0.00, 0.00, 9.00] 0.64 / 1.13
SELECT_n_of_non_aggr_func__UPPER	
0	646 (65%)
1	237 (24%)
2	71 (7.1%)
3	24 (2.4%)
4	12 (1.2%)
5	2 (0.2%)
6	4 (0.4%)
7	2 (0.2%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	[0.00, 0.00, 11.00] 0.30 / 0.88
SELECT_n_of_non_aggr_func__DAY	
0	883 (88%)
1	89 (8.9%)
2	14 (1.4%)
3	5 (0.5%)
4	7 (0.7%)
5	2 (0.2%)
SELECT_n_of_non_aggr_func__LOG	
0	854 (85%)
1	105 (11%)
2	21 (2.1%)
3	10 (1.0%)
4	5 (0.5%)
5	4 (0.4%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__LTRIM	
0	641 (64%)
1	234 (23%)
2	73 (7.3%)
3	27 (2.7%)

Query parameter or value	Descriptive statistics
4	13 (1.3%)
5	5 (0.5%)
6	5 (0.5%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__ROUND	
0	912 (91%)
1	66 (6.6%)
2	16 (1.6%)
3	6 (0.6%)
SELECT_n_of_non_aggr_func__DOW	
0	857 (86%)
1	97 (9.7%)
2	27 (2.7%)
3	7 (0.7%)
4	8 (0.8%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func__TRUNC	
0	911 (91%)
1	77 (7.7%)
2	11 (1.1%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func__FLOOR	
0	930 (93%)
1	55 (5.5%)
2	11 (1.1%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 79.0] 5.0 / 8.6
SELECT_n_of_aggr_func__MAX	[0.00, 0.00, 13.00] 0.75 / 1.78
SELECT_n_of_aggr_func__COUNT_DISTINCT	[0.00, 0.00, 13.00] 0.74 / 1.81
SELECT_n_of_aggr_func__AVG	
0	945 (95%)
1	35 (3.5%)
2	14 (1.4%)
3	3 (0.3%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 15.00] 0.93 / 2.04
SELECT_n_of_aggr_func__SUM	
0	933 (93%)
1	46 (4.6%)
2	12 (1.2%)
3	6 (0.6%)

Query parameter or value	Descriptive statistics
4	3 (0.3%)
SELECT_n_of_aggr_func__MIN	[0.00, 0.00, 12.00] 0.78 / 1.83
SELECT_n_of_all_aggr_func	[0.00, 3.00, 16.00] 3.38 / 3.10
FROM_n_of_join_paths	
1	166 (17%)
2	173 (17%)
3	166 (17%)
4	161 (16%)
5	159 (16%)
6	175 (18%)
FROM_n_of_super_joins__LEFT	
0	545 (55%)
1	304 (30%)
2	114 (11%)
3	34 (3.4%)
4	3 (0.3%)
FROM_n_of_super_joins__FULL	
0	571 (57%)
1	290 (29%)
2	117 (12%)
3	17 (1.7%)
4	5 (0.5%)
FROM_n_of_super_joins__RIGHT	
0	533 (53%)
1	327 (33%)
2	112 (11%)
3	24 (2.4%)
4	4 (0.4%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.33 / 2.53
FROM_n_of_joins__RIGHT	[0.00, 3.00, 12.00] 3.23 / 2.41
FROM_n_of_processed_rows	[5, 18,595,583, 83,754,602] 23,299,075 / 19,269,089
WHERE_n_of_predicates	[0.00, 3.00, 12.00] 3.88 / 2.93
WHERE_n_of_attribs_of_type__character_varying	
0	350 (35%)
1	265 (27%)
2	208 (21%)
3	117 (12%)
4	46 (4.6%)
5	8 (0.8%)
6	3 (0.3%)
7	3 (0.3%)
WHERE_n_of_attribs_of_type__character	
0	657 (66%)
1	243 (24%)
2	80 (8.0%)
3	16 (1.6%)

Query parameter or value	Descriptive statistics
4	4 (0.4%)
WHERE_n_of_attribs_of_type__integer	
0	368 (37%)
1	293 (29%)
2	204 (20%)
3	79 (7.9%)
4	45 (4.5%)
5	8 (0.8%)
6	3 (0.3%)
WHERE_n_of_attribs_of_type__numeric	
0	560 (56%)
1	276 (28%)
2	117 (12%)
3	37 (3.7%)
4	7 (0.7%)
5	3 (0.3%)
WHERE_n_of_attribs_of_type__date	
0	768 (77%)
1	185 (19%)
2	47 (4.7%)
WHERE_n_of_pkey_attribs	
0	528 (53%)
1	316 (32%)
2	118 (12%)
3	33 (3.3%)
4	5 (0.5%)
WHERE_n_of_connect_OR	[0.00, 2.00, 10.00] 2.25 / 2.41
WHERE_n_of_operators__greater_or_less	
0	322 (32%)
1	289 (29%)
2	181 (18%)
3	115 (12%)
4	55 (5.5%)
5	23 (2.3%)
6	14 (1.4%)
8	1 (0.1%)
WHERE_n_of_operators__between	
0	530 (53%)
1	283 (28%)
2	120 (12%)
3	46 (4.6%)
4	20 (2.0%)
5	1 (0.1%)
WHERE_n_of_operators__in	
0	527 (53%)
1	288 (29%)

Query parameter or value	Descriptive statistics
2	119 (12%)
3	52 (5.2%)
4	13 (1.3%)
5	1 (0.1%)
WHERE_n_of_operators_like	
0	774 (77%)
1	198 (20%)
2	26 (2.6%)
3	2 (0.2%)
WHERE_n_of_non_aggr_func_ROUND	
0	941 (94%)
1	57 (5.7%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_SQRT	
0	857 (86%)
1	136 (14%)
2	7 (0.7%)
WHERE_n_of_non_aggr_func_FLOOR	
0	949 (95%)
1	50 (5.0%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_LOG	
0	920 (92%)
1	74 (7.4%)
2	5 (0.5%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_DAY	
0	959 (96%)
1	41 (4.1%)
WHERE_n_of_non_aggr_func_DOW	
0	952 (95%)
1	47 (4.7%)
2	1 (0.1%)
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	42 (4.2%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	970 (97%)
1	29 (2.9%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_YEAR	
0	959 (96%)

Query parameter or value	Descriptive statistics
1	41 (4.1%)
WHERE_n_of_all_non_aggr_func	
0	580 (58%)
1	277 (28%)
2	117 (12%)
3	20 (2.0%)
4	6 (0.6%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 2.91 / 2.83
HAVING_n_of_main_predicates	
0	312 (31%)
1	81 (8.1%)
2	138 (14%)
3	165 (17%)
4	149 (15%)
5	109 (11%)
6	46 (4.6%)
HAVING_n_of_main_predicates__scalar_subquery	
0	413 (41%)
1	190 (19%)
2	204 (20%)
3	193 (19%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 18.0] 2.3 / 3.3
HAVING_n_of_processed_rows_by_subqueries	[0, 3, 459, 760, 1, 340, 073, 632] 56,821,780 / 116,536,499
ORDER_BY_n_of_columns	
0	296 (30%)
1	353 (35%)
2	223 (22%)
3	88 (8.8%)
4	19 (1.9%)
5	10 (1.0%)
6	10 (1.0%)
7	1 (0.1%)
limit	[1, 510, 997] 498 / 295
offset	[0, 0, 995] 262 / 337