

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

Scenario: 1-6jp_rjpp__g__h0-3ss_h0nss

Scale factor: 500 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, 9, 179] 16 / 24
SELECT_n_of_non_aggr_func__LTRIM	[0.00, 0.00, 11.00] 0.65 / 1.11
SELECT_n_of_non_aggr_func__UPPER	
0	610 (61%)
1	262 (26%)
2	77 (7.7%)
3	34 (3.4%)
4	4 (0.4%)
5	7 (0.7%)
6	4 (0.4%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func__YEAR	
0	820 (82%)
1	122 (12%)
2	30 (3.0%)

Query parameter or value	Descriptive statistics
3	15 (1.5%)
4	7 (0.7%)
5	2 (0.2%)
6	1 (0.1%)
7	2 (0.2%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func__DAY	
0	863 (86%)
1	111 (11%)
2	16 (1.6%)
3	8 (0.8%)
4	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func__LOG	
0	862 (86%)
1	106 (11%)
2	19 (1.9%)
3	8 (0.8%)
4	5 (0.5%)
SELECT_n_of_non_aggr_func__LOWER	
0	592 (59%)
1	274 (27%)
2	89 (8.9%)
3	26 (2.6%)
4	3 (0.3%)
5	7 (0.7%)
6	9 (0.9%)
SELECT_n_of_non_aggr_func__MONTH	[0.00, 0.00, 10.00] 0.36 / 0.96
SELECT_n_of_non_aggr_func__RTRIM	[0.00, 0.00, 11.00] 0.65 / 1.12
SELECT_n_of_non_aggr_func__SQRT	
0	791 (79%)
1	145 (15%)
2	34 (3.4%)
3	16 (1.6%)
4	11 (1.1%)
5	3 (0.3%)
SELECT_n_of_non_aggr_func__TRUNC	
0	926 (93%)
1	63 (6.3%)
2	9 (0.9%)
3	1 (0.1%)
4	1 (0.1%)
SELECT_n_of_non_aggr_func__DOW	
0	837 (84%)
1	119 (12%)
2	26 (2.6%)

Query parameter or value	Descriptive statistics
3	8 (0.8%)
4	7 (0.7%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
SELECT_n_of_non_aggr_func___SUBSTR	
0	620 (62%)
1	245 (25%)
2	84 (8.4%)
3	26 (2.6%)
4	15 (1.5%)
5	4 (0.4%)
6	3 (0.3%)
7	2 (0.2%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___ABS	
0	927 (93%)
1	63 (6.3%)
2	4 (0.4%)
3	4 (0.4%)
4	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___FLOOR	
0	923 (92%)
1	64 (6.4%)
2	11 (1.1%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	933 (93%)
1	56 (5.6%)
2	9 (0.9%)
3	1 (0.1%)
5	1 (0.1%)
SELECT_n_of_all_non_aggr_func	[0, 3, 66] 5 / 8
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 13.00] 0.91 / 2.07
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 14.00] 0.86 / 2.04
SELECT_n_of_aggr_func___AVG	
0	931 (93%)
1	43 (4.3%)
2	19 (1.9%)
3	6 (0.6%)
4	1 (0.1%)
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.97 / 2.05
SELECT_n_of_aggr_func___SUM	
0	935 (94%)
1	42 (4.2%)

Query parameter or value	Descriptive statistics
2	14 (1.4%)
3	6 (0.6%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_aggr_func__COUNT	[0.00, 0.00, 10.00] 0.74 / 1.76
SELECT_n_of_all_aggr_func	[0.0, 3.0, 16.0] 3.7 / 3.1
FROM_n_of_join_paths	
1	174 (17%)
2	143 (14%)
3	152 (15%)
4	171 (17%)
5	185 (19%)
6	175 (18%)
FROM_n_of_super_joins__FULL	
0	530 (53%)
1	303 (30%)
2	124 (12%)
3	38 (3.8%)
4	5 (0.5%)
FROM_n_of_super_joins__LEFT	
0	533 (53%)
1	321 (32%)
2	118 (12%)
3	23 (2.3%)
4	5 (0.5%)
FROM_n_of_super_joins__RIGHT	
0	545 (55%)
1	304 (30%)
2	121 (12%)
3	28 (2.8%)
4	2 (0.2%)
FROM_n_of_joins__INNER	[0.00, 3.00, 12.00] 3.40 / 2.47
FROM_n_of_joins__RIGHT	[0.00, 3.00, 13.00] 3.39 / 2.49
FROM_n_of_processed_rows	[5, 5,025,028,337, 21,710,169,482] 6,073,062,451 / 4,758,801,447
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.76 / 2.91
WHERE_n_of_attribs_of_type__integer	
0	373 (37%)
1	288 (29%)
2	187 (19%)
3	103 (10%)
4	34 (3.4%)
5	12 (1.2%)
6	3 (0.3%)
WHERE_n_of_attribs_of_type__character_varying	
0	371 (37%)

Query parameter or value	Descriptive statistics
1	282 (28%)
2	208 (21%)
3	76 (7.6%)
4	48 (4.8%)
5	13 (1.3%)
6	2 (0.2%)
WHERE_n_of_attribs_of_type__date	
0	786 (79%)
1	178 (18%)
2	34 (3.4%)
3	2 (0.2%)
WHERE_n_of_attribs_of_type__numeric	
0	564 (56%)
1	287 (29%)
2	108 (11%)
3	23 (2.3%)
4	17 (1.7%)
5	1 (0.1%)
WHERE_n_of_attribs_of_type__character	
0	649 (65%)
1	248 (25%)
2	82 (8.2%)
3	18 (1.8%)
4	2 (0.2%)
5	1 (0.1%)
WHERE_n_of_pkey_attribs	
0	526 (53%)
1	302 (30%)
2	126 (13%)
3	36 (3.6%)
4	10 (1.0%)
WHERE_n_of_connect_OR	[0.00, 1.00, 10.00] 2.13 / 2.36
WHERE_n_of_operators__greater_or_less	
0	323 (32%)
1	266 (27%)
2	197 (20%)
3	126 (13%)
4	56 (5.6%)
5	25 (2.5%)
6	7 (0.7%)
WHERE_n_of_operators__between	
0	566 (57%)
1	267 (27%)
2	112 (11%)
3	38 (3.8%)
4	16 (1.6%)

Query parameter or value	Descriptive statistics
5	1 (0.1%)
WHERE_n_of_operators_in	
0	530 (53%)
1	306 (31%)
2	109 (11%)
3	44 (4.4%)
4	9 (0.9%)
5	2 (0.2%)
WHERE_n_of_operators_like	
0	788 (79%)
1	181 (18%)
2	27 (2.7%)
3	4 (0.4%)
WHERE_n_of_non_aggr_func_LOG	
0	939 (94%)
1	59 (5.9%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func_MONTH	
0	970 (97%)
1	28 (2.8%)
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	
0	964 (96%)
1	35 (3.5%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_DOW	
0	970 (97%)
1	30 (3.0%)
WHERE_n_of_non_aggr_func_SQRT	
0	879 (88%)
1	108 (11%)
2	12 (1.2%)
3	1 (0.1%)
WHERE_n_of_non_aggr_func_ROUND	
0	956 (96%)
1	43 (4.3%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func_TRUNC	
0	957 (96%)
1	40 (4.0%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func_ABS	

Query parameter or value	Descriptive statistics
0	958 (96%)
1	41 (4.1%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func__DAY	
0	969 (97%)
1	31 (3.1%)
WHERE_n_of_all_non_aggr_func	
0	627 (63%)
1	274 (27%)
2	72 (7.2%)
3	22 (2.2%)
4	5 (0.5%)
GROUP_BY_n_of_columns	[0.00, 3.00, 13.00] 3.24 / 2.97
HAVING_n_of_main_predicates	
0	297 (30%)
1	106 (11%)
2	134 (13%)
3	187 (19%)
4	124 (12%)
5	96 (9.6%)
6	56 (5.6%)
HAVING_n_of_main_predicates__scalar_subquery	
0	434 (43%)
1	195 (20%)
2	174 (17%)
3	197 (20%)
HAVING_n_of_subqueries__scalar_subquery	[0.0, 1.0, 20.0] 2.5 / 3.6
HAVING_n_of_processed_rows_by_subqueries	0, 22,500,075, 250,181,978,900] 15,190,761,684 / 30,690,552,402
ORDER_BY_n_of_columns	
0	323 (32%)
1	309 (31%)
2	240 (24%)
3	92 (9.2%)
4	20 (2.0%)
5	6 (0.6%)
6	4 (0.4%)
7	4 (0.4%)
8	2 (0.2%)
limit	[5, 496, 1,000] 497 / 286
offset	[0, 0, 999] 246 / 326