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

Scenario: 1-6jp_ujjp__g__h0-3ss_h0nss

Scale factor: 500 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, 9, 189] 16 / 25
SELECT_n_of_non_aggr_func___ABS	
0	920 (92%)
1	68 (6.8%)
2	11 (1.1%)
3	1 (0.1%)
SELECT_n_of_non_aggr_func___LOG	
0	881 (88%)
1	97 (9.7%)
2	15 (1.5%)
3	4 (0.4%)
4	2 (0.2%)
5	1 (0.1%)
SELECT_n_of_non_aggr_func___LTRIM	
0	613 (61%)

Query parameter or value	Descriptive statistics
1	250 (25%)
2	84 (8.4%)
3	29 (2.9%)
4	8 (0.8%)
5	6 (0.6%)
6	7 (0.7%)
7	2 (0.2%)
9	1 (0.1%)
SELECT_n_of_non_aggr_func___RTRIM	
0	604 (60%)
1	278 (28%)
10	1 (0.1%)
2	81 (8.1%)
3	20 (2.0%)
4	10 (1.0%)
5	1 (0.1%)
6	3 (0.3%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___SQRT	
0	820 (82%)
1	135 (14%)
2	25 (2.5%)
3	13 (1.3%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___UPPER	[0.00, 0.00, 10.00] 0.61 / 1.11
SELECT_n_of_non_aggr_func___SUBSTR	
0	642 (64%)
1	241 (24%)
10	2 (0.2%)
2	75 (7.5%)
3	14 (1.4%)
4	11 (1.1%)
5	8 (0.8%)
6	5 (0.5%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___LOWER	
0	614 (61%)
1	246 (25%)
2	93 (9.3%)
3	23 (2.3%)
4	13 (1.3%)
5	2 (0.2%)

Query parameter or value	Descriptive statistics
6	4 (0.4%)
7	2 (0.2%)
8	3 (0.3%)
SELECT_n_of_non_aggr_func___TRUNC	
0	913 (91%)
1	76 (7.6%)
2	9 (0.9%)
3	2 (0.2%)
SELECT_n_of_non_aggr_func___DAY	
0	896 (90%)
1	80 (8.0%)
2	16 (1.6%)
3	3 (0.3%)
4	3 (0.3%)
5	1 (0.1%)
6	1 (0.1%)
SELECT_n_of_non_aggr_func___DOW	
0	869 (87%)
1	90 (9.0%)
2	23 (2.3%)
3	12 (1.2%)
4	3 (0.3%)
5	1 (0.1%)
7	1 (0.1%)
8	1 (0.1%)
SELECT_n_of_non_aggr_func___MONTH	
0	816 (82%)
1	115 (12%)
2	34 (3.4%)
3	16 (1.6%)
4	11 (1.1%)
5	5 (0.5%)
6	1 (0.1%)
7	2 (0.2%)
SELECT_n_of_non_aggr_func___YEAR	
0	842 (84%)
1	110 (11%)
2	29 (2.9%)
3	6 (0.6%)
4	5 (0.5%)
5	3 (0.3%)
6	2 (0.2%)
7	1 (0.1%)
8	2 (0.2%)
SELECT_n_of_non_aggr_func___ROUND	
0	926 (93%)

Query parameter or value	Descriptive statistics
1	62 (6.2%)
2	7 (0.7%)
3	3 (0.3%)
4	2 (0.2%)
SELECT_n_of_non_aggr_func___FLOOR	
0	929 (93%)
1	61 (6.1%)
2	7 (0.7%)
3	1 (0.1%)
4	2 (0.2%)
SELECT_n_of_all_non_aggr_func	[0.0, 3.0, 65.0] 4.7 / 7.5
SELECT_n_of_aggr_func___MAX	[0.00, 0.00, 12.00] 0.82 / 1.88
SELECT_n_of_aggr_func___MIN	[0.00, 0.00, 14.00] 0.89 / 2.10
SELECT_n_of_aggr_func___COUNT_DISTINCT	[0.00, 0.00, 12.00] 0.81 / 1.84
SELECT_n_of_aggr_func___SUM	
0	928 (93%)
1	41 (4.1%)
2	19 (1.9%)
3	10 (1.0%)
5	2 (0.2%)
SELECT_n_of_aggr_func___COUNT	[0.00, 0.00, 17.00] 0.93 / 2.13
SELECT_n_of_aggr_func___AVG	
0	937 (94%)
1	45 (4.5%)
2	14 (1.4%)
3	3 (0.3%)
4	1 (0.1%)
SELECT_n_of_all_aggr_func	[0.0, 3.0, 18.0] 3.7 / 3.3
FROM_n_of_join_paths	
1	171 (17%)
2	164 (16%)
3	160 (16%)
4	181 (18%)
5	167 (17%)
6	157 (16%)
FROM_n_of_super_joins___FULL	
0	568 (57%)
1	310 (31%)
2	93 (9.3%)
3	26 (2.6%)
4	3 (0.3%)
FROM_n_of_super_joins___RIGHT	
0	567 (57%)
1	300 (30%)
2	101 (10%)
3	29 (2.9%)

Query parameter or value	Descriptive statistics
4	3 (0.3%)
FROM_n_of_super_joins___LEFT	
0	539 (54%)
1	315 (32%)
2	118 (12%)
3	24 (2.4%)
4	4 (0.4%)
FROM_n_of_joins___INNER	[0.00, 3.00, 12.00] 3.43 / 2.55
FROM_n_of_joins___RIGHT	[0.00, 3.00, 13.00] 3.25 / 2.47
FROM_n_of_processed_rows	[5, 4,652,528,345, 21,785,169,477] 6,052,382,760 / 4,928,400,608
WHERE_n_of_predicates	[0.00, 3.00, 13.00] 3.92 / 3.01
WHERE_n_of_attribs_of_type___character	
0	639 (64%)
1	267 (27%)
2	80 (8.0%)
3	10 (1.0%)
4	4 (0.4%)
WHERE_n_of_attribs_of_type___character_varying	
0	352 (35%)
1	290 (29%)
2	202 (20%)
3	94 (9.4%)
4	41 (4.1%)
5	12 (1.2%)
6	8 (0.8%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___date	
0	778 (78%)
1	176 (18%)
2	38 (3.8%)
3	8 (0.8%)
WHERE_n_of_attribs_of_type___integer	
0	360 (36%)
1	291 (29%)
2	190 (19%)
3	88 (8.8%)
4	45 (4.5%)
5	20 (2.0%)
6	5 (0.5%)
7	1 (0.1%)
WHERE_n_of_attribs_of_type___numeric	
0	562 (56%)
1	270 (27%)
2	119 (12%)
3	38 (3.8%)

Query parameter or value	Descriptive statistics
4	9 (0.9%)
5	2 (0.2%)
WHERE_n_of_pkey_attribs	
0	522 (52%)
1	315 (32%)
2	114 (11%)
3	36 (3.6%)
4	9 (0.9%)
5	4 (0.4%)
WHERE_n_of_connect_OR	[0.00, 2.00, 12.00] 2.28 / 2.45
WHERE_n_of_operators___between	
0	505 (51%)
1	308 (31%)
2	124 (12%)
3	47 (4.7%)
4	15 (1.5%)
5	1 (0.1%)
WHERE_n_of_operators___greater_or_less	
0	330 (33%)
1	266 (27%)
2	185 (19%)
3	120 (12%)
4	63 (6.3%)
5	26 (2.6%)
6	10 (1.0%)
WHERE_n_of_operators___in	
0	526 (53%)
1	283 (28%)
2	132 (13%)
3	42 (4.2%)
4	14 (1.4%)
5	3 (0.3%)
WHERE_n_of_operators___like	
0	775 (78%)
1	190 (19%)
2	30 (3.0%)
3	4 (0.4%)
4	1 (0.1%)
WHERE_n_of_non_aggr_func___MONTH	
0	965 (97%)
1	34 (3.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___FLOOR	
0	941 (94%)
1	56 (5.6%)
2	3 (0.3%)

Query parameter or value	Descriptive statistics
WHERE_n_of_non_aggr_func___ROUND	
0	970 (97%)
1	28 (2.8%)
2	2 (0.2%)
WHERE_n_of_non_aggr_func___SQRT	
0	874 (87%)
1	116 (12%)
2	10 (1.0%)
WHERE_n_of_non_aggr_func___LOG	
0	921 (92%)
1	78 (7.8%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___DAY	
0	955 (96%)
1	44 (4.4%)
2	1 (0.1%)
WHERE_n_of_non_aggr_func___TRUNC	
0	953 (95%)
1	47 (4.7%)
WHERE_n_of_non_aggr_func___DOW	
0	963 (96%)
1	34 (3.4%)
2	3 (0.3%)
WHERE_n_of_non_aggr_func___YEAR	
0	957 (96%)
1	39 (3.9%)
2	4 (0.4%)
WHERE_n_of_non_aggr_func___ABS	
0	962 (96%)
1	37 (3.7%)
2	1 (0.1%)
WHERE_n_of_all_non_aggr_func	
0	604 (60%)
1	269 (27%)
2	94 (9.4%)
3	25 (2.5%)
4	7 (0.7%)
5	1 (0.1%)
GROUP_BY_n_of_columns	[0.00, 2.00, 12.00] 2.99 / 2.86
HAVING_n_of_main_predicates	
0	300 (30%)
1	98 (9.8%)
2	124 (12%)
3	189 (19%)
4	142 (14%)
5	86 (8.6%)

Query parameter or value	Descriptive statistics
6	61 (6.1%)
HAVING_n_of_main_predicates___scalar_subquery	
0	431 (43%)
1	179 (18%)
2	196 (20%)
3	194 (19%)
HAVING_n_of_subqueries___scalar_subquery	[0.0, 1.0, 24.0] 2.5 / 3.7
HAVING_n_of_processed_rows_by_subqueries	[0, 1,200,000,000, 246,242,033,424] 15,876,588,402 / 31,375,744,029
ORDER_BY_n_of_columns	
0	294 (29%)
1	327 (33%)
2	225 (23%)
3	110 (11%)
4	26 (2.6%)
5	9 (0.9%)
6	6 (0.6%)
7	2 (0.2%)
8	1 (0.1%)
limit	[2, 509, 1,000] 503 / 301
offset	[0, 0, 1,000] 233 / 314